

FIGURE 1

CGGCAGCCCTGACGTGATGAGCTCAACCAGCAGAGACATTCCATCCCAAGAGAGGTCTGCGTGACGCGTCCGGGAGG
 CCACCCTCAGCAAGACCACCGTACAGTTGGTGGAAGGGGTGACAGCTGCATTCTCCTGTGCCTACCACGTAACCAAA
AATGAAGGAGAACTACTGTTTACAAGCCGCCCTGGTGTGCCTGGGCATGCTGTGCCACAGCCATGCCTTTGCCCCAG
 AGCGGCGGGGGCACCTGCGGCCCTCCTTCCATGGGCACCATGAGAAGGGCAAGGAGGGGCAGGTGCTACAGCGCTCC
 AAGCGTGGCTGGGTCTGGAACCAAGTTCTTCGTGATAGAGGAGTACACCGGGCCTGACCCCGTGCTTGTGGGCAGGCT
 TCATTTCAGATATTGACTCTGGTGATGGGAACATTAAATACATTCTCTCAGGGGAAGGAGCTGGAACCATTTTTGTGA
 TTGATGACAAATCAGGGAACATTTCATGCCACCAAGACGTTGGATCGAGAAGAGAGAGAGCCAGTACACGTTGATGGCT
 CAGGCGGTGGACAGGGACACCAATCGGCCACTGGAGCCACCGTCGGAATTCATTGTCAAGGTCCAGGACATTAATGA
 CAACCCTCCGGAGTTCTTGCACGAGACCTATCATGCCAACGTGCCTGAGAGGTCCAATGTGGGAACGTCAGTAATCC
 AGGTGACAGCTTCAGATGCAGATGACCCCACTTATGGAAATAGCGCCAAGTTAGTGACAGTATCCTCGAAGGACAA
 CCTATTTTTTCGGTGGAAGCACAGACAGGTATCATCAGAACAGCCCTACCCAACATGGACAGGGAGGCCAAGGAGGA
 GTACCACGTGGTGATCCAGGCCAAGGACATGGGTGGACATATGGGCGGACTCTCAGGGACAACCAAAGTGACGATCA
 CACTGACCGATGTCAATGACAACCCACCAAAGTTTCCGCAGAGGCTATACCAGATGTCTGTGTGAGAAGCAGCCGTC
 CCTGGGGAGGAAGTAGGAAGAGTGAAAGCTAAAGATCCAGACATTGGAGAAAATGGCTTAGTCACATACAATATTGT
 TGATGGAGATGGTATGGAATCGTTTGAAATCACACGGACTATGAAACACAGGAGGGGGGTGATAAAGCTGAAAAAGC
 CTGTAGATTTTTGAAACCGAAAGAGCCTATAGCTTGAAGGTAGAGGCAGCCAACGTGCACATCGACCCGAAGTTTATC
 AGCAATGGCCCTTTCAAGGACACTGTGACCGTCAAGATCTCAGTAGAAGATGCTGATGAGCCCCCTATGTTCTTGGC
 CCAAGTTACATCCACGAAGTCCAAGAAAATGCAGCTGCTGGCACCGTGGTTGGGAGAGTGCATGCCAAAGACCCTG
 ATGCTGCCAACAGCCCGATAAGGTATTCCATCGATCGTCACACTGACCTCGACAGATTTTTCACTATTAATCCAGAG
 GATGGTTTTATTAAACTACAAAACCTCTGGATAGAGAGGAAACAGCCTGGCTCAACATCACTGTCTTTGCAGCAGA
 AATCCACAATCGGCATCAGGAAGCCCAAGTCCCAGTGGCCATTAGGGTCCTTGATGTCAACGATAATGCTCCCAAGT
 TTGCTGCCCTTATGAAGGTTTCATCTGTGAGAGTGATCAGACCAAGCCACTTTCCAACCAGCCAATTGTTACAATT
 AGTGCAGATGACAAGGATGACACGGCCAATGGACCAAGATTTATCTTCAGCCTACCCCTGAAATCATTACAATCC
 AAATTTACAGTCAGAGACAACCGAGATAACACAGCAGGCGTGTACGCGCGGTGGAGGGTTTCAGTCGGCAGAAGC
 AGGACTTGTACCTTCTGCCCATAGTGATCAGCGATGGCGGCATCCCGCCCATGAGTAGCACCAACACCCTCACCATC
 AAAGTCTGCGGGTGCGACGTGAACGGGGCACTGCTCTCCTGCAACGCAGAGGCCTACATTCTGAACGCGCGCCTGAG
 CACAGGCGCCCTGATCGCCATCCTCGCCTGCATCGTCATTCTCCTGGTCATTGTAGTATTGTTTGTGACCCTGAGAA
 GGCAAAAGAAAGAACCCTCATTGTCTTTGAGGAAGAAGATGTCCGTGAGAACATCATTACTTATGATGATGAAGGG
 GGTGGGGAAGAAGACACAGAAGCCTTTGATATTGCCACCCTCCAGAATCCTGATGGTATCAATGGATTTATCCCCG
 CAAAGACATCAAACCTGAGTATCAGTACATGCCTAGACCTGGGCTCCGGCCAGCGCCCAACAGCGTGGATGTGATG
 ACTTCATCAACACGAGAATACAGGAGGCAGACAATGACCCACGGCTCCTCCTTATGACTCCATTCAAATCTACGGT
 TATGAAGGCAGGGGCTCAGTGGCCGGGTCCCTGAGCTCCCTAGAGTCGGCCACCACAGATTTCAGACTTGGACTATGA
 TTATCTACAGAACTGGGGACCTCGTTTTTAAGAACTAGCAGATTTGTATGGTTCCAAAGACACTTTTGATGACGATT
 CTTAACAATAACGATACAAATTTGGCCTTAAGAACTGTGTCTGGCGTTCTCAAGAATCTAGAAGATGTGTAACAGGT
 ATTTTTT

FIGURE 2

AACTCAAACCTCCTCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGGGTGTTGGAGCCCTCG
GTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTCATGTATGGCAAGAGCTCTACTCGTGCGGTGCTTC
TTCTCCTTGGCATAACAGCTCACAGCTCTTTGGCCTATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCT
GTTAATGGGACAGATGCTCGGTAAAAATGCACTTTCTCCAGCTTTGCCCCGTGTGGGTGATGCTCTAACAGTGACCTG
GAATTTTCGTCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTCCAACCCATGAGTG
GGCGGTTTAAGGACCGGGTGCTTGGGATGGGAATCCTGAGCGGTACGATGCCTCCATCCTTCTCTGGAACTGCAG
TTCGACGACAATGGGACATACACCTGCCAGGTGAAGAACCCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCT
CAGCGTCGTGCACACTGTACGCTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGA
TCATAATAGTAATTGTAGTGGTCCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTCATAAAGTGGTG
GAGATAAAATCAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCTGTTTATTTAGAAGACACAGACTTAACA
ATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAAGAACCCTAGTATTTCTTGAAGTTAATGGAACTTTTCTTTG
GCTTTTCCAGTTGTGACCCGTTTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGC
CAGCACAGTGCTCCTCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTACAGAGTGTAATT
TTTTCAAGTGCTCATTAGGTTTTATAAACAAGAAGCTACATTTTTGCCCTTAAGACACTACTTACAGTGTTATGACT
TGTATACACATATATTGGTATCAAAGGGGATAAAAGCCAATTTGTCTGTTACATTTCCTTTCACGTATTTCTTTTAG
CAGCACTTCTGCTACTAAAGTTAATGTGTTTACTCTCTTCCCTTCCCACATTCTCAATTAAAAGGTGAGCTAAGCCT
CCTCGGTGTTTCTGATTAACAGTAAATCCTAAATTCAAAGTGTAAATGACATTTTTATTTTTATGTCTCTCCTTAA
CTATGAGACACATCTTGTTTTACTGAATTTCTTTCAATATTCCAGGTGATAGATTTTTGTCTG

FIGURE 3

GGCACGAGGCGATT CAGGGGAGGGAGCAACTGGAGCCTCAGGCCCTCCAGAGTAGTCTGCCTGACCACCCTGGAGCC
CACAGAAGCCCAGGACGTCTCCCGCGAGGCCTCCCCGTGTGTGGCTGAGGATGGCTGAGCAGCAGGGCCGGGAGCTT
GAGGCTGAGTGCCCCGTCTGCTGGAACCCCTTCAACAACACGTTCCATAACCCCCAAAATGCTGGATTGCTGCCACTC
CTTCTGCGTGGAATGTCTGGCCCACCTCAGCCTTGTGACTCCAGCCCGGCGCCGCTGCTGTGCCACTCTGTGCGC
AGCCCACAGTGCTGGCCTCAGGGCAGCCTGTCACTGACTTGCCACGGACACTGCCATGCTCACCTGCTCCGCCTG
GAGCCCCACCATGT CATCCTGGAAGGCCATCAGCTGTGCCTCAAGGACCAGCCCAAGAGCCGCTACTTCTGCGCCA
GCCTCGAGTCTACACGCTGGACCTTGGCCCCCAGCCTGGGGGCCAGACTGGGCGCCCCCAGACACGGCCTCTGCCA
CCGTGTCTACGCCCATCCTCATCCCCAGCCACCACTCTTTGAGGGAGTGTTTCCGCAACCCTCAGTTCCGCATCTTT
GCCTACCTGATGGCCGT CATCCTCAGTGTCCTCTGTTGCTCATATTCTCCATCTTTTGGACCAAGCAGTTCCTTTG
GGGTGTGGGGTGAAGTGCTGTTCCCAGACAAGAAACCAAACCTTTTTTCGGTTGCTGCTGGGTATGGTGACTACGGAGC
CTCATTTGGTATTGTCTTCCTTTGTAGTGTTGTTTATTTTACAATCCAGGGATTGTTTCAGGCCATGTGTTTGCTTCT
GGGAACAATTTAAAAAAAACAAAAAACGAAAAGCTTGAAGGACTGGGAGATGTGGAGCGACCTCCGGGTGTGAG
TGTGGCGTCATGGAAGGGCAGAGAAGCGGTTCTGACCACAGAGCTCCACAGCAAGTTGTGCCAAGGGCTGCACAGT
GGTATCCAGGAACCTGACTAGCCCAAATAGCAAGTTGCATTTCTCACTGGAGCTGCTTCAAAATCAGTGCATATTTT
TTTGAGTTGCTCTTTTACTATGGGTTGCTAAAAAAAAAAAAAAAAAATTGGGAAGTGAGCTTCAATTCTGTGGGTA
AATGTGTGTTTGTCTCTTTGAATGTCTTGCCACTGGTTGCAGTAAAAGTGTTCTGTATTCATTAAAAAAAAAAAA
AA

FIGURE 4

GAAGCGCGCTCCCGGGGAGGTGTTGCAGCCATGGCTACGGCAGCCGGCGCGACCTACTTTTCAGCGAGGCAGTCTGTT
CTGGTTCACAGTCATCACCCCTCAGCTTTGGCTACTACACATGGGTGTCTTCTGGCCTCAGAGTATCCCTTATCAGA
ACCTTGGGCCCCCTGGGCCCCCTTCACTCAGTACTTGGTGGACCACCATCACACCCTCCTTTGCAATGGGTATTGGCTT
GCCTGGCTGATTCATGTGGGAGAGTCCTTGTATGCCATAGCATTGTGCAAGCATAAAGGCATCACAAGTGGTCGGGC
TCAGTACTCTGGTTCCTACAGACTTTCTTCTTTGGGATAGCGTCTCTCACCATCTTGATTGCTTACAAACGGAAGC
GCCAAAAACAAACTTGAAGTTGTCTGAAAGCTTGCTCTACACTTTTACATTTCATCCTCACCCCTTTTTTTTGTGGGGT
AGAGGAGGTGCAGTAATTTACTCAGTGATCTTTCTACTTTCTAGAACTGTCCTTCAAAGCTCTTTAAGACCCCTC
GTTAGTCAGTTTCTTCTCTTATATGCTCTGGTTGAGCTTGAATAGACCAGTTGTTACTTAAGAAAGAAACAGAGAAA
GATTTTAGCTTTTCAATCCTATTTGGCAGAGGACTTCAGCTACCTTCTTACAGTCTTTGGCTGTGTTGGTACCCTCG
TGTGCTCTGAGCTAAGCCACATACTAACTGACTTTTTGGTTTGTATACCCTTGCTCCCGCCTTCTGATGAAAACAC
CTTACCCTCACAACCACCATCTTTCCTCTCCTTTCCAAAGCTCTTCCACCTTGCTGCTAAGATAAAGTGACACT
TCCACTATATGTCAATTCACACACATTTATTAGGTACCTGTGAGGTAGGATCCTATCCTCTCAAACCTTCCATTTCT
CATGCTACAGAGAAAGATAAGGAAGATGAGCAAGTGCCTGGAATGGGGCAGGCTGAGCAGTCACACAGGCATAGAGG
CACGCTGAGAACCTGGAGGGGAGACTGCAGAGTGCCTTCCCTGATGCTGCAGCCGGAAGTGATCCTTCCCTCCACCT
GGCCCCCTGGGACACTGTGCTCTGCAGTGTGCAGGGCCTGATGGCACTGCTAGATTGCTCCTTCAGCTCAGGGCCACA
GCTTAAACAGCTTTACCTTTCCCCTCAGCACCTGTCCCCTATCTTGACACAGGTGCTCTAACCATGTTTATTGAA
CAAAGGAGGGAACTGATTTCACTTTCACTTGTTTCATTATCATTTCCAATTTTTATGTGAAAATGGCACAACCCATTT
GGGGTACCTCACCCTCAAATAAAAGCCCAAGTCTACCTTTGACTGGTACCACCTTTTTTGTGGTTTCGTTGGTGAG
AAACCTTTATCTTTTTCATACCTTTCTATTCTCAATCACTTCTCCAAAAGTGTGTCTTTCCAGCTCTGATTTATTCA
AAACACAAGCATTTCTGTTTAGAGATTCTAGCCCATGGGTTATCTGGCTAGTTATTACCTCTCCTGTTCACTTAGTT
ATACTTTATTATTGCTCACAGGCTGGGGAGGCAGAATGACTCTGTCACCACTAGGAGCCATTAGGGCTTCTTCCCTG
GAGGACTGCCTGCTTGCTTTCTGGGGACACTAGCCCTCATTTCCCTTCTGTGGTACAGTGGGGCAAATTATTTGTAT
TAAGCAAACATTTATGGGAAACAACCCGCTCCCGAAAACGAGCCCCCAAGTAAAGCACAAACCTGAAAGATTATGA
ACTATGAATTGTCTCTAGTAGAGATAAATTTCTGCAAACATATCTCAGTCTTCCCTCTGTTTCTCTGGTGATTAAGA
AGTTCCTTTTTGGTAAGGAAAAGGATTTTTAACCATAGAGTTAGGCATCATGGAAATTCAAACCAGATTTCTTAATA
CCTGGTCTTCCCTCAAAGAGAAATAATAACAGTAATAGTGGTGTGGGAACAATATGGCAGATTATTGAATGAAATTG
ATTAACCTGAATAAAATGCTGTGAATTTTC

FIGURE 5

GGCACGAGGCCGCGAGCGGACTGCCCTTTCCCAAGATGGCGTCTCGAAGATAGGTTTCGAGACGGTGGATGTTGCAGCTGA
TCATGCAGTTGGGTTTCGGTGCTGCTCACACGCTGCCCCTTTTGGGGCTGCTTCAGCCAGCTCATGCTGTACGCTGAG
AGGGCTGAGGCACGCCGGAAGCCCGACATCCCAGTGCCTTACCTGTATTTTCGACATGGGGGCGAGCCGTGCTGTGCGC
TAGTTTCATGTCCTTTGGCGTGAAGCGGCGCTGGTTCGCGCTGGGGGCGCACTCCAATTGGCCATTAGCACCTACG
CCGCCTACATCGGGGGCTACGTCCACTACGGGGACTGGCTGAAGGTCCGTATGTACTCGCGCACAGTTGCCATCATC
GGCGGCTTTCTTGTTGGCCAGCGGTGCTGGGGAGCTGTACCGCCGAAACCTCGCAGCCGCTCCCTGCAGTCCAC
CGGCCAGGTGTTTCTGGGTATCTACCTCATCTGTGTGGCTACTCACTGCAGCACAGCAAGGAGGACCGGCTGGCGT
ATCTGAACCATCTCCCAGGAGGGGAGCTGATGATCCAGCTGTTCTTCGTGCTGTATGGCATCCTGGCCCTGGCCTTT
CTGTCAGGCTACTACGTGACCCTCGCTGCCCAGATCCTGGCTGTACTGCTGCCCCCTGTCATGCTGCTCATTGATGG
CAATGTTGCTTACTGGCACAACACGCGGCGTGTGAGTTCTGGAACCAGATGAAGCTCCTTGGAGAGAGTGTGGGCA
TCTTCGGAAGTCTGTCATCCTGGCCACTGATGGCTTGAGTTTTATGGCAAGAGGCTGAGATGGGCACAGGGAGCCAC
TGAGGGTCACCCTGCCTTCCTCCTTGCTGGCCCAGCTGCTGTTTATTTATGCTTTTTGGTCTGTTTGTGTTGATCTTT
TGCTTTTTTAAAATTGTTTTTTGCAGTTAAGAGGCAGCTCATTTGTCCAAATTTCTGGGCTCAGCGCTTGGGAGGGC
AGGAGCCCTGGCACTAATGCTGTACAGGTTTTTTTTCTGTTAGGAGAGCTGAGGCCAGCTGCCCCTGAGTCTCCTG
TCCCTGAGAAGGGAGTATGGCAGGGCTGGGATGCGGCTACTGAGAGTGGGAGAGTGGGAGACAGAGGAAGGAAGATG
GAGATTGGAAGTGAGCAAATGTGAAAAATTCCTCTTTGAACCTGGCAGATGCAGCTAGGCTCTGCAGTGCTGTTTGG
AGACTGTGAGAGGGAGTGTGTGTGTTGACACATGTGGATCAGGCCAGGAAGGGCACAGGGGCTGAGCACTACAGAA
GTCACATGGGTTCTCAGGGTATGCCAGGGGCGAGAAACAGTACCGGCTCTCTGTCACTCACCTTGAGAGTAGAGCAGA
CCCTGTTCTGCTCTGGGCTGTGAAGGGGTGGAGCAGGCAGTGGCCAGCTTTGCCCTTCCTGCTGTCTCTGTTTCTAG
CTCCATGGTTGGCCTGGTGGGGGTGGAGTTCCCTCCCAAACACCAGACCACACAGTCCCTCCAAAAATAAACATTTTA
TATAGACAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6

GGCCGCGGCTCGCCTTTGGCCCTTCTTATCAGGATGAAAACGCTTCTGTTTGGTGTCTGGGCCCTGCTGGCCTTGAT
CCTTTGCCCAGGGGTCCCGGAAGAGTTGTTTGAGGTTTCTATTTGGCCAAGTCAGGCCCTGGTGGAGTTTGGACAGT
CCCTAGTGTGCAACTGCAGCACTACTTGCCCAGACCCAGGACCCAGTGGAATTGAGACCTTCTTAAAGAAAACCTCAG
GTGGACAAAGGGCCTCAGTGGAAAGAGTTTCTTCTGGAGGATGTCACAGAGAATTCCATCCTGCAGTGCTTCTTCTC
TTGTGCAGGGATTCAAAGGACACAAGCCTTGGCATCACTGTGTATCAGCCACCAGAGCAAGTGATCCTGGAGCTGC
AGCCTGCCTGGGTGGCCGTGGACGAAGCCTTCACAGTGAAGTGTATGTACCCAGTGTAGCACCCCTTGGAGAGTCTC
ACCCTTGCCCTTCTCCAGGGTAACCAAGAACTGCATAGAAAAGAACTTTACGAGCTTGGCTGTGGCCTCCCAAAGAGC
TGAAGTCATCATCAGTGTGAGAGCCCAAAAGGAGAATGACAGATGCAATTCTTCCTGCCATGCAGAACTGGACTTGA
GTTTGCAAGGTGGGAGGCTCTTTCAAGGCAGCTCACCCATCAGAATAGTCCGGATCTTTGAATTCTCTCAGAGTCCC
CACATCTGGGTCTCTTCCCTTTTGGAGGCTGGGATGGCGGAGACTGTGAGCTGCGAGGTGGCTAGGGTGTTCAGC
CAAAGAAGTTATGTTCCACATGTTTCTGGAAGACCAAGAGCTGAGCTCCTCCCTTTCTGGGAGGGGGACACAGCAT
GGGCAATGCTACCATTTCGGACCATGGAGGCTGGTGATCAGGAAGTGTCTTGCTTTGCATCTCTGGGTGCAATGGAA
CAGAAGACAAGAAAGCTAGTGATAGCTACAGCTTCCCTCCACCAATCCTGGAGCTAAAAGAATCATACCCATTGGC
AGGGACCGACATTAATGTGACCTGCTCAGGGCATGTATTAACATCACCCAGCCCTACTCTTCGGCTTCAGGGAGCCC
CAGACCTCCCTGCTGGGGAGCCTGCCTGGCTTCTACTTACTGCCAGGGAGGAAGATGATGGCXGAAATTTCTCCTGC
GAGGCCTCTTTGGTGGTGCAGGGTCAGCGTTGATGAAAACCACTGTGATCCAGCTCCATATCCTAAAGCCACAGTT
AGAGGAATCCAGTTGCCCTGGCAAACAGACCTGGCTGGAAGGGATGGAACACACGCTCGCCTGCGTCCCAAAGGGAA
ACCCAGCTCCAGCCTTGGTGTGTACCTGGAATGGGGTGGTCTTTGACCTTGAAGTGCCACAGAAGGCAACCTAGAAC
CACACTGGAACCTACCGCTACACAGCCACTAACCAGCTGGGCTCTGTCAGCAAAGACATTGCTGT

FIGURE 7

CCACGCGTCCGTTCTGAGGTGCATTCTTTTTTTGATGAGAGGCATCTCTAGGTACCATCCCTGACCTGGTCCTC**ATG**
 CTGCCGAGGCTGTTGCTGTTGATCTGTGCTCCACTCTGTGAACCTGCCGAGCTGTTTTTGATAGCCAGCCCCCTCCCA
 TCCCACAGAGGGGAGCCCAGTGACCCTGACGTGTAAGATGCCCTTTCTACAGAGTTCAGATGCCAGTTCAGTTCT
 GCTTTTTTCAGAGACACCCGGGCCTTGGGCCCAGGCTGGAGCAGCTCCCCCAAGCTCCAGATCGCTGCCATGTGGAAA
 GAAGACACAGGGTCATACTGGTGCGAGGCACAGACAATGGCGTCCAAAGTCTTGAGGAGCAGGAGATCCCAGATAAA
 TGTGCACATCCCGGTGTCTCGCCCAATCCTCATGCTCAGGGCTCCCAGGGCCCAGGCTGCAGTGGAGGATGTGCTGG
 AGCTTCACTGTGAGGCCCTGAGAGGCTCTCCTCCAATCCTGTACTGGTTTTATCACGAGGATATCACCTGGGGAGC
 AGGTCGGCCCCCTCTGGAGGAGGAGCCTCCTTCAACCTTTCCCTGACTGAAGAACATTCTGGAACTACTCCTGTGA
 GGCCAACAATGGCCTGGGGGGCCAGCGCAGTGAGGCGGTGACACTCAACTTCACAGTGCCTACTGGGGCCAGAAGCA
 ATCATCTTACCTCAGGAGTCATTGAGGGGCTGCTCAGCACCTTGGTCCAGCCACCGTGGCCTTATTATTTTGCTAC
 GGCTCAAAAGAAAAATAGGAAGACGTTTCCAGCCAGGGATCCACTCAGGAGCCTTCCAGCCTTACCCCAAGAGTTCAC
 CTACCTCAACTCACCTACCCAGGGCAGCTACAGCCTATATATGAAAATGTGAATGTTGTAAGTGGGGATGAGGTTT
 ATTCAGTGGCGTACTATAACCAGCCGGAGCAGGAATCAGTAGCAGCAGAAACCTGGGGACACATATGGAGGACAAG
 GTTTCCTTAGACATCTATTCCAGGCTGAGGAAAGCAAACATTACAGATGTGGACTATGAAGATGCTATG**TAAG**GTTA
 TGGAAGATTCTGCTCTTTGAAAACCATCCATGACCCCAAGCCTCAGGCCTGATATGTTCTTCAGAGATCCTGGGGCA
 TTAGCTTTCCAGTATACCTCTTCTGGATGCCATTCTCCATGGCACTATTCTTCATCTACTGTGAAGTGAAGTTGGC
 GCAGCCCTGAAGAACTACCTAGGAGAACTAATAGACACAGGAGTGACAGGGACTTTGTTATCAGAACCAGATTCTT
 GCCGGCTCCTTTGAAAACAGGTCATATTGTGCTCTTCTGTTTACAAGAGGAAACAAGATGGAATAAAAGAAATTGGG
 ATCTTGGGTTGGAGGGACAGTGAAGCTTAGAGCACATGAACTCAAGGTTAGTGACTCTGCAGGACTTCACAGAGAGA
 GCTGTGCCCATCATTCAGTCCAAGTGCTTTCTCTGCCAGACAGCACAGAACTCCAGCCCCGCTACTTACATGGATC
 ATCGAGTTTCCACCTAAAATATGATTCTATTTATTTTGAGTCACTGTTACCAAATTAGAACTAAAACAAAGTTACAT
 AAAAAGTTATTGTGACTCCACTTAATTTTAGTGACGTATTTTGTATATATAGGCCAACCTATACCACATCCAAAAT
 TATGTATCTATTACAGCCCCTAGAAGCTTTATAAATACAGTGTGTCTTCTTTTATTCACAAAATTTTTGAAATCGTG
 GTAATATGGTTTGAAACCTGTATCTTAATTATTTTTTTTTTAAATTGAGACAGGGTCTCACTCTGTCACTCAATCTG
 GAATGCAGTGGCACAATCTTGCTCACTGCAACGCCTGCCTCTCAGGCTCAAGCAAACCTCTCACCTCAGCCTGCTG
 AGTAGCTGGGACTACAGGCACATGCCACCAAACCTTGCCATTTTTTGTCTTACGTAGAGACAAGATTTACCGTTTTT
 GCCCAGGCTGGTCTCAAACCTCTGGGCTCAAGCAATGTATTGAATTTTAAATAACCAGGCACTCACTCTTATGAAT
 TAATAAACATTTGGAGGTATATAAAGTAAAAAGTTAAAGTCTTTCCTGTAAGTTAACACAAATGTTAACTATTGTTA
 AAAACTTTACAGGTAGCTCTCTAGATATTTTTCTATTTTGTATGTATACTTATGCATACATGTAAGTATATAAACA
 TTTAGAAGTGACCTATCTAACAACTATTATGAAATACTTTCAAATCTGTAAATAGATCTATTATACTATTTTAAA
 AGTCTCTATAGTAGTGTGTATATAGATAAATCATAACTTTTTCTTTTTTATTGTAGTAAATATGCACAACATAA
 AATTGATCATTTTAACCATTTTTAAGTGTACAATTCAGTGGCATTAAGTACTATCATAATATATTTAATCCTTCTC
 ATCACTGGTGGACATTAAGGAGACTCTCAAAAAATTCATATTATAAAAAACAAAGTTCAAACAAATGTCTTTGTACTA
 GCATATTATGGCACTCCTGCTGGATTATCTGAAGGATAAATTTGTAAATCTAGTATTGCTAGATTATGCATATTAAA
 TATTCTTGTTAAATAGTCAAAAAAAAAAAAAAAAAA

FIGURE 8

CTCAATCAGCTTTATGCAGAGAAGAAGCTTACTGAGCTCACTGCTGGTGCTGGTGCTAGGCAAGTGCTGCTTTGGCAA
 TCTGGGCTGACCTGGCTTGTCTCCTCAGAACTCCTTCTCCAACCCCTGGAGCAGGCTTCC**ATG**CTGCTGTGGGCGTCC
 TTGCTGGCCTTTTGCTCCAGTCTGTGGACAATCTGCAGCTGCACACAAACCTGTGATTTCCGTCCATCCTCCATGGAC
 CACATTCTTCAAAGGAGAGAGAGTGACTCTGACTTGCAATGGATTTTCAGTTCTATGCAACAGAGAAAACAACATGGT
 ATCATCGGCACTACTGGGGAGAAAAGTTGACCCTGACCCAGGAAACACCCCTCGAGGTTTCGGGAATCTGGACTGTAC
 AGATGCCAGGCCCGGGGCTCCCCACGAAGTAACCCTGTGCGCTTGCTCTTTTCTTCAGACTCCTTAATCCTGCAGGC
 ACCATATTCTGTGTTTGAAGGTGACACATTGGTTCCTGAGATGCCACAGAAGAAGGAAAGAGAAATTGACTGCTGTGA
 AATATACTTGGAATGGAAACATTCTTTCCATTTCTAATAAAAGCTGGGATCTTCTTATCCCACAAGCAAGTTCAAAT
 AACAAATGGCAATTATCGATGCATTGGATATGGAGATGAGAATGATGATTTTAGATCAAATTTCAAAATAATTAAAA
 TCAAGAACTATTTCCACATCCAGAGCTGAAAGCTACAGACTCTCAGCCTACAGAGGGGAATTCTGTAAACCTGAGCT
 GTGAAACACAGCTTCTCCTCCAGAGCGGTGAGACACCCCACTTCACTTCAACTTCTTCAGAGATGGCGAGGTATCCTG
 TCAGACTGGAGCACGTACCCGGAACCTCAGCTCCCAACCGTCTGGAGAGAAAACCTCAGGATCCTATTGGTGTGGTGC
 TGAAACAGTGAGGGGTAAACATCCACAAGCACAGTCCCTCGCTACAGATCCATGTGCAGCGGATCCCTGTGTCTGGGG
 TGCTCCTGGAGACCCAGCCCTCAGGGGGCCAGGCTGTTGAAGGGGAGATGCTGGTCCCTTGTCTGCTCCGTGGCTGAA
 GGCACAGGGGATACCAATTCTCCTGGCACCGAGAGGACATGCAGGAGAGTCTGGGGAGGAAAACCTCAGCGTTCCCT
 GAGAGCAGAGCTGGAGCTCCCTGCCATCAGACAGAGCCATGCAGGGGGATACTACTGTACAGCAGACAACAGCTACG
 GCCCTGTCCAGAGCATGGTGTGAATGTCACTGTGAGAGAGACCCAGGCAACAGAGATGGCCTTGTGCGCCGCGGGA
 GCCACTGGAGGGCTGCTCAGTGCTCTTCTCCTGGCTGTGGCCCTGCTGTTTCACTGCTGGCGTCCGAGGAAGTCAGG
 AGTTGGTTTTCTTGGGAGACGAAACCAGGCTCCCTCCCGCTCCAGGCCAGGAGAGTCTCCCATTCATCTGCCCTG
 CCCAGGTGGAGCTTCAGTCGTTGTATGTTGATGTACACCCAAAAAGGGAGATTTGGTATACTCTGAGATCCAGACT
 ACTCAGCTGGGAGAAGAAGAGGAAGCTAATACCTCCAGGACACTTCTAGAGGATAAGGATGTCTCAGTTGTCTACTC
 TGAGGTAAAGACACAACACCCAGATAAATCAGCTGGAAAAGATCAGCTCTAAGGATGAAGAAAAGTT**TA**AGAGAATGAAA
 AGTTACGGGAACGTCCTACTCATGTGATTTCTCCCTTGTCCAAAGTCCAGGCCAGTGCAGTCTTGTGGGCACCTG
 GAATGATCAACTCATTCCAGCTTTCTAATTCTTCTCATGCATATGCATTCACTCCCAGGAATACTCATTCGTCTACT
 CTGATGTTGGGATGGAATGGCCTCTGAAAGACTTCACTAAAATGACCAGGATCCACAGTTAAGAGAAGACCCTGTAG
 TATTTGCTGTGGGCCTGACCTAATGCATTCCCTAGGGTCTGCTTTAGAGAAGGGGGATAAAGAGAGAGAAGGACTGT
 TATGAAAAACAGAAGCACAAATTTTGGTGAATTGGGATTTGCAGAGATGAAAAAGACTGGGTGACCTGGATCTCTGC
 TTAATACATCTACAACCATTGTCTCACTGGAGACTCACTTGCATCAGTTTGTTTAACTGTGAGTGGCTGCACAGGCA
 CTGTGCAACAATGAAAAGCCCTTCACTTCTGCCTGCACAGCTTACACTGTGAGGATTCAGTTGCAGATTAAAGAA
 CCCATCTGGAATGGTTTACAGAGAGAGGAATTTAAAGAGGACATCAGAAGAGCTGGAGATGCAAGCTCTAGGCTGC
 GCTTCCAAAAGCAAATGATAATTATGTTAATGTCATTAGTGACAAAGATTTGCAACATTAGAGAAAAAGAGACACAAA
 TATAAAATTA AAAACTTAAGTACCAACTCTCCAAAATAAATTTGAACTTAAATATTAGTATAAACTCATAATAAA
 CTCTGCCTTTTAAAAAAGATAAAATATTTCTACGTCTGTTCACTGAAATAATTACCAACCCCTTAGCAATAAGCACT
 CCTTGCAGAGAGGTTTTATTCTCTAAATACCATTCCCTTCTCAAAGGAAATAAGGTTGCTTTTCTTGTAGGAACTGT
 GTCTTTGAGTTACTAATTAGTTTATATGAGAATAATCTTGCATAAATGAAGAAGGAATAAAAGAAATAGGAAGCC
 ACAAATTTGTATGGATATTTTCATGATACACCTACTGGTTAAATAATTGACAAAAACCAGCAGCCAAATATTAGAGGT
 CTCCTGATGGAAGTGTAATAACCACTACAAATTATCCATGCCCCAAGTGTTAAAACTGAATCCATTCAAGTCTTT
 CTAAGTGAATACTTGTTTTATAGAAAATGCATGGAGAAAAGGAATTTGTTTAAATAACATTATGGGATTGCAACCAG
 CAAAAATAAACTGAGAAAAAGTTCTATAGGGCAATCACCTGGCTTCTATAACAAATAAATGGGAAAAAAATGAAA
 TAAAAAGAAGAGAGGGAGGAAGAAAGGGAGAGAGAAGAAAAGAAAAATGAAGAAAAGTAATTAGAATATTTTCAACA
 TAAAGAAAAGACGAATATTTAAGGTGACAGATATCCCAACTACGCTGATTTGATCTTTACAAATTATATGAGTGTAT
 GAATTTGTACATGTATCACCCCCAAAAAAGAGAAAAAAGAAAAATAGAAGACATATAAATTAAATGAGACGAGACA
 TGTGACCAAAAAGGAATGTGTGGGTCTTGTTTGGATCCTGACTCAAATTAAGAAAAAATAAACTACCTACGAAATA
 CTAAGAAAAATTTGTATACTAATATTAAGAAATGTTGTGTGTTTTGGATATAAGTGATAGTTTATTGTAGTGTGT
 TTTTATAAAAAGCAAAAGGATATTCACCTTTCAGCGCTTATACTGAAGTATTAGATTAAAGCTTATTAACGTA

FIGURE 9

GCCGAGCTGAGCGGATCCTCACATGACTGTGATCCGATTCTTTCCAGCGGCTTCTGCAACCAAGCGGGTCTTACCCC
CGGTCCCTCCGCGTCTCCAGTCCTCGCACCTGGAACCCCAACGTCCCCGAGAGTCCCCGAATCCCCGCTCCCAGGCTA
CCTAAGAGGATGAGCGGTGCTCCGACGGCCGGGGCAGCCCTGATGCTCTGCGCCGCCACCGCCGTGCTACTGAGCGC
TCAGGGCGGACCCGTGCAGTCCAAGTCGCCGCGCTTTGCGTCCTGGGACGAGATGAATGTCTGGCGCACGGACTCC
TGCAGCTCGGCCAGGGGCTGCGCGAACACGCGGAGCGCACCCGCAGTCAGCTGAGCGCGCTGGAGCGGCGCCTGAGC
GCGTGCGGGTCCGCTGTGAGGGAACCGAGGGGTCCACCGACCTCCCGTTAGCCCCCTGAGAGCCGGGTGGACCCTGA
GGTCCTTCACAGCCTGCAGACACAACCTCAAGGCTCAGAACAGCAGGATCCAGCAACTCTTCCACAAGGTGGCCAGC
AGCAGCGGCACCTGGAGAAGCAGCACCTGCGAATTCAGCATCTGCAAAGCCAGTTTGGCCTCCTGGACCACAAGCAC
CTAGACCATGAGGTGGCCAAGCCTGCCCGAAGAAAGAGGCTGCCCGAGATGGCCAGCCAGTTGACCCGGCTCACAA
TGTCAGCCGCTGCACCGGCTGCCAGGGATTGCCAGGAGCTGTTCCAGGTTGGGGAGAGGCAGAGTGGACTATTTG
AAATCCAGCCTCAGGGGTCTCCGCCATTTTTGGTGAAGTGAAGATGACCTCAGATGGAGGCTGGACAGTAATTCAG
AGGCGCCACGATGGCTCAGTGGACTTCAACCGGCCCTGGGAAGCCTACAAGGCGGGGTTTGGGGATCCCCACGGCGA
GTTCTGGCTGGGTCTGGAGAAGGTGCATAGCATCACGGGGACCGCAACAGCCGCTGGCCGTGCAGCTGCGGGACT
GGGATGGCAACGCCGAGTTGCTGCAGTTCTCCGTGCACCTGGGTGGCGAGGACACGGCCTATAGCCTGCAGCTCACT
GCACCCGTGGCCGGCCAGCTGGGCGCCACCACCGTCCCACCCAGCGGCCTCTCCGTACCCCTTCTCCACTTGGGACCA
GGATCACGACCTCCGCAGGGACAAGAACTGCGCCAAGAGCCTCTCTGGAGGCTGGTGGTTTGGCACCTGCAGCCATT
CCAACCTCAACGGCCAGTACTTCCGCTCCATCCCACAGCAGCGGCAGAAGCTTAAGAAGGGGAATCTTCTGGAAGACC
TGGCGGGGCCGCTACTACCCGCTGCAGGCCACCACCATGTTGATCCAGCCCATGGCAGCAGAGGCAGCCTCCTAGCG
TCCTGGCTGGGCCTGGTCCCAGGCCCACGAAAGACGGTGACTCTTGGCTCTGCCCGAGGATGTGGCCGTTCCCTGCC
TGGGCAGGGGCTCCAAGGAGGGGCCATCTGGAACCTTGTGGACAGAGAAGAAGACCACGACTGGAGAAGCCCCCTTT
CTGAGTGCAGGGGGGCTGCATGCGTTGCCTCCTGAGATCGAGGCTGCAGGATATGCTCAGACTCTAGAGGCGTGGAC
CAAGGGGCATGGAGCTTCACTCCTTGCTGGCCAGGGAGTTGGGGACTCAGAGGGACCACTTGGGGCCAGCCAGACTG
GCCTCAATGGCGGACTCAGTCACATTGACTGACGGGGACCAGGGCTTGTGTGGGTGAGAGCGCCCTCATGGTGCTG
GTGCTGTTGTGTGTAGGTCCCCTGGGGACACAAGCAGGCGCCAATGGTATCTGGGCGGAGCTCACAGAGTTCTTGGA
ATAAAAGCAACCTCAGAACAC

FIGURE 10

GATGTGCTCCTTGGAGCTGGTGTGCAGTGTCTGACTGTAAGATCAAGTCCAAACCTGTTTTGGAATTGAGGAACT
TCTCTTTTGATCTCAGCCCTTGGTGGTCCAGGTCTTCATGCTGCTGTGGGTGATATTACTGGTCCTGGCTCCTGTCA
GTGGACAGTTTGCAAGGACACCCAGGCCCATTTTTCCTCCAGCCTCCATGGACCACAGTCTTCCAAGGAGAGAGA
GTGACCCTCACTTGCAAGGGATTTTCGCTTCTACTCACCACAGAAAACAAAATGGTACCATCGGTACCTTGGGAAAGA
AATACTAAGAGAAACCCAGACAATATCCTTGAGGTTTCAGGAATCTGGAGAGTACAGATGCCAGGCCAGGGCTCCC
CTCTCAGTAGCCCTGTGCACTTGGATTTTTCTTCAGAGATGGGATTTTCCTCATGCTGCCAGGCTAATGTTGAACTC
CTGGGCTCAAGTGATCTGCTCACCTTAGGCCTCTCAAAGCGCTGGGATTACAGCTTCGCTGATCCTGCAAGCTCCACT
TTCTGTGTTTGAAGGAGACTCTGTGGTTCTGAGGTGCCGGGCAAAGGCGGAAGTAACACTGAATAATACTATTTACA
AGAATGATAATGTCCTGGCATTTCCTTAATAAAAGAACTGACTTCCAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 11A

CGCCCCGCGCGCTGCAGCCCCATCTCTAGCGGCAGCCCAGGCGCGGAGGGAGCGAGTCCGCCCCGAGGTAGGTCCAG
GACGGGCGCACAGCAGCAGCCGAGGCTGGCCGGGAGAGGGAGGAAGAGG**AT**GGCAGGGCCACGCCCCAGCCCATGGG
CCAGGCTGCTCCTGGCAGCCTTGATCAGCGTCAGCCTCTCTGGGACCTTGGCAAACCGCTGCAAGAAGGCCCCAGTG
AAGAGCTGCACGGAGTGTGTCCGTGTGGATAAGGACTGCGCCTACTGCACAGACGAGATGTTTCAGGGACCGGCGCTG
CAACACCCAGGCGGAGCTGCTGGCCGCGGGCTGCCAGCGGGAGAGCATCGTGGTCATGGAGAGCAGCTTCCAAATCA
CAGAGGAGACCCAGATTGACACCACCCTGCGGCGCAGCCAGATGTCCCCCAAGGCCTGCGGGTCCGTCTGCGGCCC
GGTGAGGAGCGGCATTTTGGAGCTGGAGGTGTTTGGAGCCACTGGAGAGCCCCGTGGACCTGTACATCCTCATGGACTT
CTCCAACTCCATGTCCGATGATCTGGACAACCTCAAGAAGATGGGGCAGAACCTGGCTCGGGTCTTGAGCCAGCTCA
CCAGCGACTACACTATTGGATTTGGCAAGTTTGTGGACAAGTCAGCGTCCCGCAGACGGACATGAGGCCTGAGAAG
CTGAAGGAGCCCTGGCCCCAACAGTGACCCCCCTTCTCCTTCAAGAACGTCATCAGCCTGACAGAAGATGTGGATGA
GTTCCGGAATAAACTGCAGGGAGAGCGGATCTCAGGCAACCTGGATGCTCCTGAGGGCGGCTTCGATGCCATCCTGC
AGACAGCTGTGTGCACGAGGGACATTGGCTGGCGCCCGGACAGCACCCACCTGCTGGTCTTCTCCACCGAGTCAGCC
TTCCACTATGAGGCTGATGGCGCCAACGTGCTGGCTGGCATCATGAGCCGCAACGATGAACGGTGCCACCTGGACAC
CACGGGCACCTACACCCAGTACAGGACACAGGACTACCCGTGCGTGCCACCCCTGGTGCGCCTGCTCGCCAAGCACA
ACATCATCCCCATCTTTGCTGTCACTCACTCCTATAGCTACTACAGAGAAGCTTACACACTATTTCCTGTCTCC
TCACTGGGGGTGCTGCAGGAGGACTCGTCCAACATCGTGGAGCTGCTGGAGGAGGCCTTCAATCGGATCCGCTCCAA
CCTGGACATCCGGGCCCTAGACAGCCCCCGAGGCCTTCGGACAGAGGTCACCTCCAAGATGTTCAGAAGACGAGGA
CTGGGTCTTTTACATCCGGCGGGGGGAAGTGGGTATATACCAGGTGCAGCTGCGGGCCCTTGAGCACGTGGATGGG
ACGCACGTGTGCCAGCTGCCGGAGGACCAGAAGGGCAACATCCATCTGAAACCTTCCTTCTCCGACGGCCTCAAGAT
GGACGCGGGCATCATCTGTGATGTGTGCACCTGCGAGCTGCAAAAAGAGGTGCGGTGAGCTCGCTGCAGCTTCAACG
GAGACTTCGTGTGCGGACAGTGTGTGTGCAGCGAGGGCTGGAGTGGCCAGACCTGCAACTGCTCCACCGGCTCTCTG
AGTGACATTTCAGCCCTGCCTGCGGGAGGGCGAGGACAAGCCGTGCTCCGGCCGTGGGGAGTGCCAGTGCGGGCACTG
TGTGTGTACGGCGAAGGCCGCTACGAGGGTCAGTTCTGCGAGTATGACAACCTCCAGTGTCCTCCGCACTTCCGGGT
TCCTCTGCAATGACCGAGGACGCTGCTCCATGGGCCAGTGTGTGTGTGAGCCTGGTTGGACAGGCCCAAGCTGTGAC
TGTCCTCTCAGCAATGCCACCTGCATCGACAGCAATGGGGGCATCTGTAATGGACGTGGCCACTGTGAGTGTGGCCG
CTGCCACTGCCACCAGCAGTCGCTCTACACGGACACCATCTGCGAGATCACTACTCGGCGATCCACCCGGGCCTCT
CGGAGGACCTACGCTCCTGCGTGCAGTGCCAGGCGTGCGGACCCGGCGAGAAGAAGGGGCGACAGTGTGAGGAATGC
AACTTCAAGGTCAAGATGGTGGACGAGCTTAAGAGAGCCGAGGAGGTGGTGGTGCCTGCTCCTTCCGGGACGAGGA
TGACGACTGCACCTACAGCTACACCATGGAAGGTGACGGCGCCCCCTGGGCCCAACAGCACTGTCTGGTGCACAAGA
AGAAGGACTGCCCTCCGGGCTCCTTCTGGTGGCTCATCCCCCTGCTCCTCCTCCTGCGGCTCCTGGCCCTGCTA
CTGCTGCTATGCTGGAAGTACTGTGCCTGCTGCAAGGCCCTGCCCTGGCACTTCTCCCGTGTGCAACCGAGGTACAT
GGTGGGCTTTAAGGAAGACCATACATGCTGCGGGAGAACCTGATGGCCTCTGACCACTTGGACACGCCATGCTGC
GCAGCGGGAACCTCAAGGGCCGTGACGTGGTCCGCTGGAAGGTACCAACAACATGCAGCGGCCTGGCTTTGCCACT
CATGCCGCCAGCATCAACCCACAGAGCTGGTGCCCTACGGGCTGTCTTGCGCCTGGCCCCGCTTTGCACCGAGAA
CCTGCTGAAGCCTGACACTCGGGAGTGCGCCAGCTGCGCCAGGAGGTGGAGGAGAACCTGAACGAGGTCTACAGGC
AGATCTCCGGTGTACACAAGCTCCAGCAGACCAAGTTCCGGCAGCAGCCCAATGCCGGGAAAAGCAAGACCACACC
ATTGTGGACACAGTGTGATGGCGCCCCGCTCGGCCAAGCCGGCCCTGCTGAAGCTTACAGAGAAGCAGGTGGAACA
GAGGGCCTTCCACGACCTCAAGGTGGCCCCCGGCTACTACACCTCACTGCAGACCAGGACGCCCGGGCATGGTGG
AGTTCCAGGAGGGCGTGGAGCTGGTGGACGTACGGGTGCCCTCTTTATCCGGCCTGAGGATGACGACGAGAAGCAG
CTGCTGGTGGAGGCCATCGACGTGCCCGCAGGCACTGCCACCCTCGGCCGCCGCTGGTAAACATCACCATCATCAA
GGAGCAAGCCAGAGACGTGGTGTCTTTGAGCAGCCTGAGTTCTCGGTGAGCCGCGGGGACCAGGTGGCCCGCATCC
CTGTCTATCCGGCGTGTCTGGACGGCGGGAAGTCCCAGGTCTCCTACCGCACACAGGATGGCACCGCGCAGGGCAAC
CGGGACTACATCCCCGTGGAGGGTGAGCTGCTGTTCAGCCTGGGGAGGCCTGGAAAGAGCTGCAGGTGAAGCTCCT
GGAGCTGCAAGAAGTTGACTCCCTCCTGCGGGGCCCGCCAGGTCCGCGGTTTCCACGTCCAGCTCAGCAACCCTAAGT
TTGGGGCCCACCTGGGCCAGCCCCACTCCACCACCATCATCATCAGGGACCCAGATGAACTGGACCGGAGCTTACAG
AGTCAGATGTTGTATCACAGCCACCCCTCACGGCGACCTGGGCGCCCCGAGAACCCCAATGCTAAGGCCGCTGG
GTCCAGGAAGATCCATTTCAACTGGCTGCCCCCTTCTGGCAAGCCAATGGGGTACAGGGTAAAGTACTGGATTACAGG
GTGACTCCGAATCCGAAGCCCACCTGCTCGACAGCAAGGTGCCCTCAGTGGAGCTCACCAACCTGTACCCGTATTGC
GACTATGAGATGAAGGTGTGCGCTACGGGGCTCAGGGCGAGGGACCCTACAGCTCCCTGGTGTCTGCGCACCCA
CCAGGAAGTGGCGAGCCAGGGCGTCTGGCCTTCAATGTCTCTCTCCACGGTGACCCAGCTGAGCTGGGCTG
AGCCGGCTGAGACCAACGGTGAGATCACAGCCTACGAGGTCTGCTATGGCCTGGTCAACGATGACAACCGACCTATT
GGGCCATGAAGAAAGTGTGGTTGACAACCCTAAGAACCGGATGCTGCTATTGAGAACCTTCGGGAG

FIGURE 11B

TCCCAGCCCTACCGCTACACGGTGAAGGCGCGCAACGGGGCCGGCTGGGGGCCTGAGCGGGAGGCCATCATCAACCT
GGCCACCCAGCCCCAAGAGGGCCCATGTCCATCCCCATCATCCCTGACATCCCTATCGTGGACGCCAGAGCGGGGAGG
ACTACGACAGCTTCCTTATGTACAGCGATGACGTTCTACGCTCTCCATCGGGCAGCCAGAGGCCAGCGTCTCCGAT
GACACTGAGCACCTGGTGAATGGCCGGATGGACTTTGCCTTCCCAGGAGCACCACCTCCCTGCACAGGATGACCAC
GACCAGTGCTGCTGCCTATGGCACCCACCTGAGCCACACGTGCCCCACCGCGTGCTAAGCACATCCTCCACCCTCA
CACGGGACTACAACCTCACTGACCCGCTCAGAACAACCTCACACTCGACCACACTGCCGAGGGACTACTCCACCCTCACC
TCCGTCTCCTCCCACGACTCTCGCCTGACTGCTGGTGTGCCCCGACACGCCCCACCCGCTGGTGTCTCTGCCCTGGG
GCCCACATCTCTCAGAGTGAGCTGGCAGGAGCCGCGGTGCGAGCGGCCGCTGCAGGGCTACAGTGTGGAGTACCAGC
TGCTGAACGGCGGTGAGCTGCATCGGCTCAACATCCCCAACCCCTGCCAGACCTCGGTGGTGGTGAAGACCTCCTG
CCCAACCACCTCCTACGTGTTCGCGTGCGGGCCAGAGCCAGGAAGGCTGGGGCCGAGAGCGTGAGGGTGTATCAC
CATTGAATCCCAGGTGCACCCGCAGAGCCCACTGTGTCCCCTGCCAGGCTCCGCCTTCACTTTGAGCACTCCCAGTG
CCCCAGGCCCCGCTGGTGTTCCTGAGCCCTGAGCCAGACTCGCTGCAGCTGAGCTGGGAGCGGCCACGGAGGCCCAAT
GGGGATATCGTCGGCTACCTGGTGACCTGTGAGATGGCCCAAGGAGGAGGGCCAGCCACCGCATTCGGGGTGGATGG
AGACAGCCCCGAGAGCCGGCTGACCGTGCCGGGGCCTCAGCGAGAACGTGCCCTACAAGTTCAAGGTGCAGGCCAGGA
CCACTGAGGGCTTCGGGCCAGAGCGCGAGGGCATCATCACCATAGAGTCCCAGGATGGAGGACCCTTCCCGCAGCTG
GGCAGCCGTGCCGGGCTCTTCCAGCACCCGCTGCAAAGCGAGTACAGCAGCATCACCACCACCCACACCAGCGCCAC
CGAGCCCTTCTAGTGGATGGGCCGACCCTGGGGGGCCAGCACCTGGAGGCAGGCGGCTCCCTCACCCGGCATGTGA
CCCAGGAGTTTGTGAGCCGGACACTGACCACCAGCGGAACCCCTTAGCACCCACATGGACCAACAGTTCTTCCAACT
TGACCGCACCCTGCCCCACCCCGCCATGTCCCCTAGGCGTCTCTCCGACTCCTCTCCCGGAGCCTCCTCAGCTAC
TCCATCCTTGCACCCCTGGGGGGCCAGCCACCCGCATGCACAGAGCAGGGGCTAGGTGTCTCCTGGGAGGCATGAA
GGGGGCAAGGTCCGTCTCTGTGGGCCCCAACCTATTTGTAACCAAAGAGCTGGGAGCAGCACAAAGGACCCAGCCTT
TGTCTGCACTTAATAAATGGTTTTGCTACTG

FIGURE 12A

GCCTTCAACTACCATCCCACCACCTGCTGAGGAGAAAAATTCTTCAAGACTCAGAGCACACAGCCAGCACCAGAGGC
 CCC**ATG**ACCCCTGGACAGACCAGGGGAGGGGGCCACCATGCTGAAGACATTCACTGTTTTGCTCTTTTGCTATTCGGAT
 GAGTCTGGGTATGACATCGATAGTGATGGACCTCAACCGAGTTGTGGATAGAGTCCAACCTACCCCAAGGCCCTT
 GGGAGAACATCACGCTTTGGTGCCGAAGCCCCCTCTCGGATATCAAGCAAGTTCTTGCTGCTGAAGGATAAGACACAA
 ATGACCTGGATCCGCCCTTCCCACAAGACCTTCCAAGTTTCACTTCTTATAGGTGCCCTTACTGAGTCCAATGCAGG
 TCTTTACCGGTGCTGCTACTGGAAGGAGACAGGCTGGTCAAAGCCCAGTAAAGTTCTAGAGTTGGAGGCACCAGGCC
 AACTGCCCAAGCCCATCTTCTGGATTACAGGCTGAGACCCCCGCTCTTCTGCGGTGTAATGTTAACATCCTCTGCCAT
 GGCTGGCTGCAGGATTTGGTATTGCTGCTGTTTAAAGAGGGATATGCAGAGCCTGTGGATTACCAAGTCCCAACTGG
 GACAATGGCCATATTCTCCATTGACAACCTGACACCTGAGGATGAAGGGGTTTACATCTGCCGCACTCATATCCAGA
 TGCTCCCCACCCTGTGGTCAGAGCCCAGCAACCCCTGAAGCTGGTTGTAGCAGGACTCTACCCCAAACCAACTTTG
 ACAGCCCATCCTGGGCCCATCATGGCACCTGGAGAAAGCCTGAATCTCAGGTGCCAAGGGCCAATCTATGGAATGAC
 CTTTGCTCTAATGAGGGTTGAAGACTTGGAGAAGTCTTTTTACCACAAGAAGACAATAAAAAATGAGGCAAATTTCT
 TCTTCCAGTCTTTGAAGATCCAAGATACTGGACATTACCTCTGTTTTTACTATGACGCATCATATAGAGGTTCACTC
 CTTAGTGATGTCCTGAAAATCTGGGTAACCTGACACTTCCCCAAGACCTGGCTACTTGCTCGGCCCAGTGCTGTGGT
 CCAAATGGGTGAGAATGTGAGCTACGGTGTGAGGACCACTGGATGGAGTGGGTCTTGCACTCTATAAGAAAGAG
 AAGACAAACCACCTTCAATTTTTTGGATGCCACCAGCATCGATGACAACACATCATTCTTCTCAACAATGTAACCTAC
 AGTGATACTGGCATCTATAGCTGCCACTATCTTCTCACCTGGAAGACCTCCATTAGGATGCCATCACACAACACTGT
 GGAGCTTATGGTTGTAGATAAGCCCCCAAACCTCCCTGTCAGCTTGGCCAAGCACTGTGTTCAAGCTAGGAAAGG
 CCATCACCCCTTCAGTGCCGAGTATCTCATCCAGTACTGGAATTTTCTCTGGAATGGGAAGAAAGAGAAACATTCCAA
 AGATTCTCAGTAAACGGAGACTTCATCATCAGTAATGTTGACGGGAAAGGCACAGGGACCTACAGTTGCAGCTATCG
 CGTAGAGACACATCCTAACATGTGGTCACATCGCAGTGAGCCCTGAAGCTGATGGGGCCAGCAGGCTATCTCACCT
 GGAATTACGTTCTGAATGAAGCTATCAGGTTGTCTCTAATCATGCAGCTTGTGTCCTTGCTGTTGGTAGTGCTGTGG
 ATAAGGTGGAAGTGTGCGGAGACTCAGAATCAGAGAAGCCTGGTTGCTGGGAACAGCTCAAGGGGTACCATGCTCTT
 CATAGTCACGGCCCTTCTCTGCTGTGGACTGTGCAATGGGGTATTGATAGAAGAGACTGAAATAGTCATGCCAACCC
 CTAAGCCTGAGCTGTGGGCAGAGACCAACTTTCCTCTGGCCCCGTGGAAGAACTTAACCTCTGGTGCAGAAGCCCT
 TCTGGCTCAACTAAGGAGTTTGTGTTGCTGAAGGATGGGACCGGGTGGATCGCCACTCGCCCGCCCTCAGAGCAGGT
 CCGGGCTGCCCTTCCCCCTTGGCGCCCTGACCCAGAGCCACACCGGGAGCTACCACTGCCATTCTATGGGAGGAGATGG
 CTGTATCGGAGCCAGTGAGGCACCTGAGCTGGTGGGGACAGACATCCTCCCCAAACCTGTCAATTTCTGCTTCCCCC
 ACAATCCGGGGCCAGGAACATAACTCCGGTGCAAAGGATGGCTGGCAGGCATGGGGTTTGCTCTGTATAAGGAGGG
 AGAGCAAGAACCTGTCCAGCAACTTGGTGCTGTTGGAAGAGAAGCCTTCTTTACAATCCAGAGAATGGAGGATAAAG
 ACGAAGGCAATTACAGCTGCCGCACTCACACTGAAAAATCCCCCTTCAAGTGGTCTGAGCCCAGTGAGCCGCTGGAG
 CTTGTCATAAAAGAAATGTACCCCTAAGCCCTTCTTCAAGACTGGGCCAGCCCTGTGGTCACCCCTGGTGCCGAGT
 GACTTTCAATTGCTCCACCCCCCACCAGCATATAGCTTTTATTCTTTACAAAGATGGAAGTGAAATAGCATCCAGTG
 ACAGGTCTGGGCAAGTCCGGGGGCCAGTGACGCTCACTTTCTAATCATTTCCGTGGGCATTTGGTGATGGAGGGAAT
 TACAGCTGCCGATATTATGACTTTTCTATCTGGTCTGAGCCCAGCGACCCTGTGGAGCTCGTGGTGACAGAATTCTA
 CCCCCAAACCCACTCTCCTGGCACAGCCAGGTCTGTGGTGTTCCTGGGAAGAGTGTGATCCTGCGCTGCCAAGGGA
 CTTTCCAGGGCATGAGGTTCCGCCCTCTTGACAGGAGGGAGCCCATGTTCCCTTACAGTTTCCGAGTGCTCAGGGAAC
 TCAGCTGACTTCCCTTCTCCACACTGTTGGAGCAGAGGACTCTGGGAACATAGCTGTATCTACTATGAGACAACCAT
 GTCAAACAGGGGGTCAATATCTCAGTATGCCCCCTTATGATCTGGGTGACTGACACATTCCCTAAGCCATGGTTGTTTG
 CTGAGCCCAGTTCTGTGGTTCCCATGGGGCAGAATGTTACTCTCTGGTGCCGAGGGCCGGTCCATGGAGTAGGATAC
 ATTCTGCACAAAGAAGGAGAAGCCACTTCAATGCAGCTCTGGGGATCCACCAGTAATGACGGGGCATTCCCCATCAC
 CAATATATCTGGTACTAGCATGGGGCGTTACAGCTGCTGCTACCACCCTGACTGGACCAGTTCTATCAAGATACAAC
 CTAGCAACACCCCTGGAACCTTAGTCACAGGCTTACTCCCCAAACCCAGCCTATTAGCCCAGCCTGGTCCCATGGTG
 GCCCCTGCGGAAAATATGACTCTTCAAGTGTCAAGGGGAAGTCCAGACTCAACATTTGTGCTGTTGAAGGAGGGGGC
 TCAGGAGCCTTTAGAGCAACAGAGGCCAAGTGGGTACAGGGCTGACTTCTGGATGCCAGCAGTGAGAGGTGAAGACT
 CTGGGATCTATAGCTGTGTTTATTATTTGGACTCTACTCCCTTTGCAGCTTCAAATCACAGTGAAGTCCCTGGAGATC
 TGGGTGACTGATAAGCCCCCTAAACCTCTCTGTGACGCTGGCCAGCACCATGTTCAAGTTAGGGAAGGACATCAC
 CCTTCAGTGCCGAGGACCCCTGCCAGGTGTGAATTTGTTCTAGAACATGATGGAGAAGAAGCACCTCAGCAGTTTTT
 CAGAGGATGGAGACTTTGTCTATCAACAACGTAGAAGGAAAGGCATTGGAAACTACAGTGCAGTACCSCCTCCAG
 GCCTACCCTGATATCTGGTCAGAGCCTAGTGATCCCCCTGAGCTGGTGGGGGCAGCAGGGCCTGTTGCTCAGGAGTG
 CACTGTAGGGAACATTGTCCGAAGTAGCCTAATCGTGGTGGTTGTTGTAGCCTTGGGGGTAGTGCTAGCCATAGAGT
 GGAAGAAGTGGCTCGACTGCGAACCAGAGGCTCAGAGACAGACGGAAGAGACCAGACCATTGCCCTTG

FIGURE 12B

AAGAGTGTAACCAAGAAGGAGAACCAGGCACCCCTGCCAATTCTCCTTCATCAACCTCTCAGAGAATCTCTGTGGAA
CTGCCCCGTTCCAATATAATAATCTCCTCCTTTACAAGAGCTTTCCTCTCCTCTCTTGCTCTCAGAGACCTATAAA
TCCAACCAGTTACCCTGCAAGTCAGCCCCATCTGCTGTTCCCTTGGTCTCTAATCACCTGAGCTGGGTAAAGGGGATT
CTGGGAGTTGAGAGCTCTGCCAGGGTGAGATGTTTCCTGAAGAGAGGTTCCCCACCCCTGTAACCTCCTCACTGTACT
GATTTACTGGCGCATGAAATTCTATTAAAAATGCATTCTTCTGAATAAAAAGAGTATTCACTATTTAACTTCAATT

CCGGAGCGCGGCGGAGCGGTGGCGCGGCAGAGGCGCGGCTCCAGCTTCGGCTCCGGCTCGGGCTCGGGCTCGGGCTCCGGCTCCGGCTCCGGCTCCAGCTCGGGTGGCGGTGGCGGGAGCGGGACCAGGTGGAGGCGGCGGCGGCAGAGGAGTGGGAGCAGCGGCCCTAGCGGCTTGCGGGGGGACATGCGGACCGACGGCCCTGGATAGGCGGAAGGAGTGGAGGCCCTGGTGCCCGGCTTGGTGCTGAGTATGCAGCAAGAGTGCACCGGGTGAAGAAGCAAAGACTCGGTTGATTGCTCTGGGTGTGGCTGGGTGTGAGACTAGAGCCCTGGATGGCCCCCTGAGCCAGCCAGGAGGACGATGGTGCCCCCTGTGCTGCACTGGTGATGCTTGGTTTTGGTGGCAGGCGCCCATGGTGACAGCAAACCTGTCTTCATTAAAGTCCCTGAGGACAGACTGGGCTGTCAGGAGGGGTAGCCTCCTTCGTGTGCCAAGCTACAGGAGAACCCAAGCCGCGCATCACATGGAAGAAGAAGGGGAAGAAAGTCAGCTCCCAGCGCTTCGAGGTCAATTGAGTTTGATGATGGGGCAGGGTCAGTGCTTCGGATCCAGCCATTGCGGGTGCAGCGAGATGAAGCCATCTATGAGTGTACAGCTACTAACAGCCTGGGTGAGATCAACACATGTAGCCAAAGCTCTCAGTGCTCGAAGAGGAACAGCTGCCCCCTGGGTTCCCTTCCATCGACATGGGGCCCTCAGCTGAAGTGGTGGAGAAGGCACGCACCACTGCTATGTGCCCAGCGAAATCCAGACCTGAGATTTCTTGGTTC AAGGACTTCCTTCCTGTAGACCTGCCACGAGCAACGGCCGATCAAGCAGCTGCGTTCAGGTGCTTGCAGATAGA GAGCAGTGAGGAATCCGACCAAGGCAAGTACGAGTGTGTGGCGACCAACTCGGCAGGCACACGTTACTCAGCCCCCTGCGAACCTGTATGTGCGAGTGCGCCGCGTGGCTCCTCGTTTCTCCATCCCTCCCAGCAGCCAGGAGGTGATGCCAGGC GGCAGCGTGAACCTGACATGCGTGGCAGTGGGTGCACCCATGCCCTACGTGAAGTGGATGATGGGGGCCGAGGAGCTCACAAGGAGGATGAGATGCCAGTTGGCCGCAACGCTCTGGAGCTCAGCAATGTCGTACGCTCTGCCAACCTACACCT GTGTGGCCATCTCCTCGCTGGGCATGATCAGGAGCCAGCCAGGTCACAGTGAAGACTCTTCCAAGCCCTCCGATTGATCTTGTGGTGACAGAGACAACCTGCCACCAGTGTCAACCTCAGCTGGGACTCTGGGAACTCGGAGCCTGTAACCTA CTATGGCATCCAGTACC GCGCAGCGGGCAGGAGGGCCCCCTTTCAGGAGGTGGATGGTGTGGCCACCACCCGCTACAGCATTTGGCGGCCCTCAGCCCCCTTCTCGGAATATGCCTTCCGCGTGTGGCGGTGAACAGCATCGGGCGAGGGCCGCCC AGCGAGGCGAGTGGCGGCACGCGAGGAGAACAGGCGCCCTCCAGCCCACCGCGCCGCGTGCAGGCACGCATGCTGAGCGCCAGCACCATGCTGGTGAGTGGGAGCCTCCGAGGAGCCCAACGGCCTGGTGGCGGGGATACCGGCTCTACTATACTCCGGACTCCCGCCGCCCCCGAAGCCTGGCACAAGCACAACCCAGCAGCGGGGCTCCTACGACCGTGGGCAGC CTGCTGCCTGGCATCACCTACAGCCTGCGCGTGTCTGCCTTACC GCGCGTGGGCGATGGCCCTCCCAGCCCCACCATCCAGGTCAAGACGCAGCAGGGAGTGCCTGCCAGCCCCGCGACTTCCAGGCCGAGGTGGAGTCGGACACCAGGATCCAGCTCTCGTGGCTGCTGCCCCCTCAGGAGCGGATCATCATGTATGAAGTGGTGTACTGGGCGGCAGAGGACGAAGACCAACAGCACAAAGGTCAACCTTCGACCCAAACCTCCTCTACACACTAGAGGACCTGAAGCCTGACACACTCTACCGCTTCCAGTGGCTGCAGCTCGGATATGGGGTGGGCGCTTCTCACCCCCACCATTGAGGCCCGCAGCCAGTCCACCCCTCCCGCCCTCCCCAGAAGGTGATGTGTGTGAGCATGGGCTCCACCAGCTCCGGTAAAGTTGGGTCCCGCCGCTGCCGACAGCCGCAACGGCGTTATCACCCAGTACTCCGTGGCCCCAGAGGCGGTGGACGGCGAGGACCGCGGGCGGCA TGTGGTGGATGGCATCAGCCGTGAGCACTCCAGCTGGGACCTGGTGGGCCTGGAGAAGTGGACGGAGTACCGGGTGTGGGTGCGGGCACAACAGACGTGGGGCCCCGCGCCCCGAGAGCAGCCCCGGTGTGGTGCGCACCGATGAGGACGTGCCAGCGGGGCTCCGCGGAAGGTGGAGGTGGAGGCCACTGAACCTCCACTGCTGTGCATGTCTACTGGAAGCTGCCTGTCC CAGCAAGCAGCATGGCCAGATCCGCGGCTACCAAGTCACTTACCTAGTCGGGCTGGAGATGGCAGGACCCCGTGAGACTCCCATCATCCAAGACGTCATGCTAGCCGAGGCCCAGTGGCGGCCAGTGAGGAGTCCGAGGACTATGAAACCACTATCAGCGGCTGACCCCGGAGACCACCTACTCCGTTACTGTTGCTGCCTATACCACCAAGGGGGATGGTGCCCGCAGCAAGGCCAAAATTGTCACTACAACAGGTGCAGTCCCAGGCCGGCCACCATGATGATCAGCACCACGGCCATGAACACTGCGCTGCTCCAGTGGCACCCACCAAGGAAGTGCCTGGCGAGCTGCTGGGCTACCGGCTGCAGTACTGCCGGGCGACGAGGCGCGGCCCAACACCATAGATTTCCGGAAGGATGACCAGCACTTACAGTCAACCGGCTGCACAAGGGGACCACCTA CATCTTCGGCTTGTGTGCCAAGAAGCCGGCTGCCTTGGGTGAGGAGTTCGAGAAGGAGATCAGGACCCCGAGGACCTGCCCAGCGGCTTCCCCCAAACCTGCATGTGACAGGACTGACCACGCTTACCACAGAAGTGGCCTGGGACCCGCCA GTGCTGGCGGAGAGGAACGGGCGCATCATCAGCTACACCGTGGTGTTCGAGACATCAACAGCCAACAGGAGCTGCA GAACATCACGACAGACACCCGCTTTACCTTACTGGCCTCAAGCCAGACACCACTTACGACATCAAGGTCCGCGCATGGACCAGCAAAAGGCTCTGGCCCACTCAGCCCCAGCATCCAGTCCCAGGACCATGCCGGTGGAGCAAGTGTTCGCAAG AACTTCCGGTGGCGGCTGCAATGAAGACGCTGTGTGCTCAGCTGGGAGGTTCCCGACTCCTATAAGTACAGCTGTGCCCTTTAAGATTCTGTACAATGGGCAGAGTGTGGAGGTGGACGGGCACTCGATGCGGAAGCTGATCGCAGACCTGC AGCCCAACACAGAGTACTCGTTTTGTGCTGATGAACCGTGGCAGCAGCGCAGGGGGCCTGCAGCACCTGGTGTCCATC CGCACAGCCCCGACCTCCTGCCTCACAAGCCGCTGCCTGCCTCTGCCTACATAGAGGACGGCCGCTTCGATCTCTC CATGCCCATGTGCAAGACCCCTCGCTTGTGAGGTGGTTCTACATTGTGTGGTACCCATTGACCGTGTGGGCGGGA GCATGCTGACGCCAAGGTGGAGCACACCCGAGGAAGTGGAGCTGGACGAGCTTCTAGAAGCCATCGAGCAAGGCGGA GAGGAGCAGCGGCGCGCGCGGCGGCGGAGCAAGCTGAAGCCATGTGGCTGCTCAACTGATGTGCTCCCGA GACCTTTACCTTGGGGGCAAGAAGAACTACCGGGGCTTCTACAACCGGCCCTGTCTCCGCACTTGAGCT

FIGURE 13B

ACCAGTGTCTTGTGCTTGCCTCCTTGAAGGAACCCATGGACCAGAAGCGCTATGCCTCCAGCCCCCTACTCGGATGAG
 ATCGTGGTCCAGGTGACACCAGCCCAGCAGCAGGAGGAGCCGGAGATGCTGTGGGTGACGGGTCCCGTGTGGCAGT
 CATCCTCATCATCCTCATTGTTCATCGCCATCCTCTTGTTCAAAGGAAAAGGACCCACTCTCCGTCTCTAAGGATG
 AGCAGTGCATCGGACTGAAGGACTCCTTGCTGGCCCACTCCTCTGACCCTGTGGAGATGCGGAGGCTCAACTACCAG
 ACCCCAGGTATGCGAGACCACCCACCCATCCCCATCACCGACCTGGCGGACAACATCGAGCGCCTCAAAGCCAAACGA
 TGGCCTCAAGTTCTCCCAGGAGTATGAGTCCATCGACCCTGGACAGCAGTTTACGTGGGAGAATTCAAACCTGGAGG
 TGAACAAGCCCCAAGAACCGCTATGCGAATGTTCATCGCCTACGACCACTCTCGAGTCATCCTTACCTCTATCGATGGC
 GTCCCCGGGAGTGACTACATCAATGCCAACTACATCGATGGCTACCGCAAGCAGAATGCCACATCGCCACGCAGGG
 CCCCCTGCCGAGACCATGGGCGATTTCTGGAGAAATGGTGTGGGAACAGCGCACGGCCACTGTGGTTCATGATGACAC
 GGCTGGAGGAGAAGTCCCGGTAAAATGTGATCAGTACTGGCCAGCCCGTGGCACCGAGACCTGTGGCCTTATTTCAG
 GTGACCCTGTTGGACACAGTGGAGCTGGCCACATACACTGTGCGCACCTTCGCACTCCACAAGAGTGGCTCCAGTGA
 GAAGCGTGAGCTGCGTCAGTTTCAGTTCATGGCCTGGCCAGACCATGGAGTTCTGAGTACCCAACCTCCCATCCTGG
 CCTTCCTACGACGGGTCAAGGCCTGCAACCCCTAGACGCAGGGCCCATGGTGGTGCATGTCAGCGCGGGCGTGGGC
 CGCACCGGCTGCTTCATCGTGATTGATGCCATGTTGGAGCGGATGAAGCAGAGAAGACGGTGGACATCTATGGCCA
 CGTGACCTGCATGCCATCACAGAGGAACCTACATGGTGCAGACGGAGGACCAAGTACGTGTTTCATCCATGAGGCGCTGC
 TGGAGGCTGCCACGTGCGGCCACACAGAGGTGCCTGCCCCGAACCTGTATGCCACATCCAGAAGCTGGGCCAAGTG
 CCTCCAGGGGAGAGTGTGACCGCCATGGAGCTCGAGTTCAAGTTGCTGGCCAGCTCCAAGGCCACACGTCCCGCTT
 CATCAGCGCCAACCTGCCCTGCAACAAGTTCAAGAACCGGCTGGTGAACATCATGCCCTACGAATTGACCCGTGTGT
 GTCTGCAGCCCATCCGTGGTGTGGAGGGCTCTGACTACATCAATGCCAGCTTCTGGATGGTTATAGACAGCAGAAG
 GCCTACATAGCTACACAGGGGCTCTGGCAGAGAGCACCGAGGACTTCTGGCGCATGCTATGGGAGCACAATTCAC
 CATCATCGTCATGCTGACCAAGCTTCGGGAGATGGGCAGGGAGAAATGCCACCAAGTACCTGGCCAGCAGAGCGCTGC
 CTCGCTACCAGTACTTTGTTGTTGACCCGATGGCTGAGTACAACATGCCCCAGTATATCCTGCGTGAGTTCAAGGTC
 ACGGATGCCCCGGGATGGGCAGTCAAGGACAATCCGGCAGTTCCAGTTTACAGACTGGCCAGAGCAGGGCGTGGCCAA
 GACAGGCGAGGGATTCAATTGACTTCATCGGGCAGGTGCATAAGACCAAGGAGCAGTTTGGACAGGATGGGCCTATCA
 CGGTGCATGTCAGTGTGGCGTGGGCCGCACCGGGGTGTTTCATCACTCTGAGCATCGTCTGGAGCGCATGCGCTAT
 GAGGGCGTGGTCGACATGTTTCAGACCGTGAAGACCCTGCGTACACAGCGTCTGCCATGGTGCAGACAGAGGACCA
 GTATCAGCTGTGCTACCGTGGGCCCTGGAGTACCTCGGCAGCTTTGACCACTATGCAACGTAACCTACCGCTCCCCCT
 CTCTCCGCCACCCCCGCGTGGGGCTCCGGAGGGGACCCAGCTCCTCTGAGCCATACCGACCATCGTCCAGCCCTC
 CTACGAGATGCTGTCACTGGCAGAGCACAGCCACGGGGATCACAGCGTTTCAGGAACGTTGCCACACCAATCAGA
 GAGCCTAGAACATCCCTGGGCAAGTGGATGGCCAGCAGGCAGGCACCTGTGGCCCTTCTGTCCACCAGACCCACCTG
 GAGCCCGCTTCAAGCTCTCTGTTGCGCTCCCGCATTTCTCATGCTTCTTCTCATGGGGTGGGGTGGGGCAAAGCCT
 CCTTTTAAATACATTAAGTGGGGTAGACTGAGGGATTTTAGCCTCTTCCCTCTGATTTTTCTTTTCGCGAATCCGTA
 TCTGCAGAATGGGCCACTGTAGGGGTGGGGTTTATTTTGTGTTTTTTTTTTTTTTTTTTTGTATGACTTCTGCTGA
 AGGACAGAACATTGCCTTCCCTCGTGCAGAGCTGGGGCTGCCAGCCTGAGCGGAGGCTCGGCCGTGGGCCGGGAGGCA
 GTGCTGATCCGGCTGCTCCTCCAGCCCTTCAGACGAGATCCTGTTTCAGCTAAATGCAGGGAAACTCAATGTTTTTT
 TAAGTTTGTGTTTTCCCTTTAAAGCCTTTTTTTAGGCCACATTGACAGTGGTGGGCGGGGAGAAGATAGGGAACACTC
 ATCCCTGGTCTGCTATCCAGTGTGTGTTTAAACATTACAGCCAGAACCACAGATGTGTCTGGGAGAGCCTGGCAA
 GGCATTCTCATCACCATCGTGTGTTGCAAAGGTTAAACAAAAAACAAAAAACCAAAAAATAAAAAACAAAAAAC
 AAAAAACCAAAAAAAGAGTCAAGCCCTTGGCTTCTGCTTCAAACCCTCAAGAGGGGAAGCAACTCCG
 TGTGCTGGGGTTCCCGAGGGAGCTGCTGGCTGACCTGGGCCACAGAGCCTGGCTTTGGTCCCCAGCATTCAGTA
 TGGTGTGGTGTGTTGTTAGGCTGTGGGGTCTGGCTGTGTGGCCAAGGTGAATAGCACAGGTTAGGGTGTGTGCCACACC
 CCATGCACCTCAGGGCCAAGCGGGGGCGTGGCTGGCCTTTCAGGTCCAGGCCAGTGGGCCTGGTAGCACATGTCTGT
 CCTCAGAGCAGGGGCCAGATGATTTTCTCCTGGTTTGACGTGTTTTCAAAGCCCCGATAATCGTCTTTTCCA
 CTCCAAGATGCCCTCATAAACCAATGTGGCAAGACTACTGGACTTCTATCAATGGTACTCTAATCAGTCCTTATTAT
 CCCAGCTTGCTGAGGGGAGGGAGAGCGCCTCTTCTCTGGGCAGCGCTATCTAGATAGGTAAGTGGGGGCGGGGAA
 GGGTGCATAGCTGTTTTAGCTGAGGGACGTGGTGGCGACGTCCCCAAACCTAGCTAGGCTAAGTCAAGATCAACATT
 CCAGGGTTGGTAATGTTGGATGATGAAACATTCAATTTTACCTTGTGGATGCTAGTGTGAGAGTTCACTGTTGTA
 CACAGTCTGTTTTCTATTTGTTAAGAAAACTACAGCATATTGCATAATTCTTGATGGTAATAAATTTGAATAATC
 AGATTTCT

FIGURE 14

GGAGAGGTGCGGGCCGAATCCGAGCCGAGCGAGAGGAATCCGGCAGTAGAGAGCGGACTCCAGCCGGCGGACCCTGC
AGCCCTCGCCTGGGACAGCGGCGCGCTGGGCAGGCGCCCAAGAGAGCATCGAGCAGCGGAACCCGCGAAGCCGGCCC
GCAGCCGCGACCCGCGCAGCCTGCCGCTCTCCCGCCGCGGTCCGGGCAGCATGAGGCGCGCGGCGCTCTGGCTCTG
GCTGTGCGCGCTGGCGCTGAGCCTGCAGCTGGCCCTGCCGCAAATTGTGGCTACTAATTTGCCCCCTGAAGATCAAG
ATGGCTCTGGGGATGACTCTGACAACCTTCTCCGGCTCAGGTGCAGGTGCTTTGCAAGATATCACCTTGTCACAGCAG
ACCCCTCCACTTGGAAGGACACGCAGCTCCTGACGGCTATTCCACGTCTCCAGAACCCACCGGCTGGAGGCTAC
AGCTGCCTCCACCTCCACCCTGCCGGCTGGAGAGGGGGCCCAAGGAGGGAGAGGCTGTAGTCCTGCCAGAAGTGGAGC
CTGGCCTCACCGCCCGGGAGCAGGAGGCCACCCCGACCCAGGGAGACCACACAGCTCCCGACCACTCATCAGGCC
TCAACGACCACAGCCACCACGGCCCAGGAGCCCGCCACCTCCACCCCCACAGGGACATGCAGCCTGGCCACCATGA
GACCTCAACCCCTGCAGGACCCAGCCAAGCTGACCTT.CACACTCCCCACACAGAGGATGGAGGTCTTCTGCCACCG
AGAGGGCTGCTGAGGATGGAGCCTCCAGTCAGCTCCCAGCAGCAGAGGGCTCTGGGGAGCAGGACTTCACCTTTGAA
ACCTCGGGGGAGAATACGGCTGTAGTGGCCGTGGAGCCTGACCGCCGGAACCAGTCCCCAGTGGATCAGGGGGCCAC
GGGGGCCTCACAGGGCCTCCTGGACAGGAAAGAGGTGCTGGGAGGGGTCAATTGCCGGAGGCCTCGTGGGGCTCATCT
TTGCTGTGTGCCTGGTGGGTTTCATGCTGTACCGCATGAAGAAGAAGGACGAAGGCAGCTACTCCTTGAGGAGCCG
AAACAAGCCAACGGCGGGGCTACCAGAAGCCCACCAAACAGGAGGAATTCTATGCCTGACGCGGGAGCCATGCGCC
CCCTCCGCCCTGCCACTCACTAGGCCCCCACTTGCCCTCTTCTTGAAGAACTGCAGGCCCTGGCCTCCCCTGCCACC
AGGCCACCTCCCCAGCATTCCAGCCCCCTCTGGTCGCTCCTGCCCACGGAGTCGTGGGTGTGCTGGGAGCTCCACTCT
GCTTCTCTGACTTCTGCCTGGAGACTTAGGGCACCAGGGTTTCTCGCATAGGACCTTTCCACCACAGCCAGCACCT
GGCATCGCACCATTCTGACTCGGTTTCTCCAACTGAAGCAGCCTCTCCCCAGGTCCAGCTCTGGAGGGGAGGGGGA
TCCGACTGCTTTGGACCTAAATGGCCTCATGTGGCTGGAAGATCTGCGGGTGGGGCTTGGGGCTCACACACCTGTAG
CACTTACTGGTAGGACCAAGCATCTTGGGGGGGTGGCCGCTGAGTGGCAGGGACAGGAGTCACTTTGTTTCGTGGGG
AGGTCTAATCTAGATATCGACTTGTTTTTGCACATGTTTCCTCTAGTTCTTTGTTTCATAGCCCAGTAGACCTTGTTA
CTTCTGAGGTAAGTTAAGTAAGTTGATTTCGGTATCCCCCATCTTGCTTCCCTAATCTATGGTCGGGAGACAGCATC
AGGGTTAAGAAGACTTTTTTTTTTTTTTTTAACTAGGAGAACCAAATCTGGAAGCCAAAATGTAGGCTTAGTTTG
TGTGTTGTCTCTTGAGTTTGTCGCTCATGTGTGCAACAGGGTATGGACTATCTGTCTGGTGGCCCCGTTTCTGGTGG
TCTGTTGGCAGGCTGGCCAGTCCAGGCTGCCGTGGGGCCGCCGCTCTTTCAAGCAGTCGTGCCTGTGTCCATGCGC
TCAGGGCCATGCTGAGGCCTGGGCCGTGCCACGTTGGAGAAGCCCGTGTGAGAAGTGAATGCTGGGACTCAGCCTT
CAGACAGAGAGGACTGTAGGGAGGGCGGCAGGGGCTGGAGATCCTCCTGCAGACCACXCCCGTCTGCCTGTGCGC
CGTCTCCAGGGGCTGCTTCCTCCTGGAAATTGACGAGGGGTGCTTTGGGCAGAGCTGGCTCTGAGCGCCTCCATCCA
AGGCCAGGTTCTCCGTTAGCTCCTGTGGCCCCACCCTGGGCCCTGGGCTGGAATCAGGAATATTTTCCAAAGAGTGA
TAGTCTTTTGCTTTTGGCAAACTCTACTTAATCCAATGGGTTTTTCCCTGTACAGTAGATTTTCCAAATGTAATAA
ACTTTAATATAAAGT

FIGURE 15

GCTCCGGCCAGCCGCGGTCCAGAGCGCGGAGGTTTCGGGGAGCTCCGCCAGGCTGCTGGTACCTGCGTCCGCCCGGC
GAGCAGGACAGGCTGCTTTGGTTTGTGACCTCCAGGCAGGACGGCCATCCTCTCCAGAATGAAGATCTTCTTGCCAG
TGCTGCTGGCTGCCCTTCTGGGTGTGGAGCGAGCCAGCTCGCTGATGTGCTTCTCCTGCTTGAACCAGAAGAGCAAT
CTGTACTGCCTGAAGCCGACCATCTGCTCCGACCAGGACAACCTACTGCGTGACTGTGTCTGCTAGTGCCGGCATTGG
GAATCTCGTGACATTTGGCCACAGCCTGAGCAAGACCTGTTCCCCGGCCTGCCCCATCCCAGAAGGCGTCAATGTTG
GTGTGGCTTCCATGGGCATCAGCTGCTGCCAGAGCTTTCTGTGCAATTTAGTGCGGCCGATGGCGGGCTGCGGGCA
AGCGTCACCCTGCTGGGTGCCGGGCTGCTGCTGAGCCTGCTGCCGGCCCTGCTGCGGTTTGGCCCCTGACCGCCAG
ACCCTGTCCCCCGATCCCCAGCTCAGGAAGGAAAGCCCAGCCCTTTCTGGATCCCACAGTGTATGGGAGCCCCTGA
CTCCTCACGTGCCTGATCTGTGCCCTTGGTCCCAGGTCAGGCCCACCCCTGCACCTCCACCTGCCCCAGCCCCTGC
CTCTGCCCAAGTGGGCCAGCTGCCCTCACTTCTGGGGTGGATGATGTGACCTTCCTTGGGGGACTGCGGAAGGGACG
AGGGTTCCCTGGAGTCTTACGGTCCAACATCAGACCAAGTCCCATGGACATGCTGACAGGGTCCCCAGGGAGACCGT
GTCAGTAGGGATGTGTGCCTGGCTGTGTACGTGGGTGTGCAGTGCACGTGAGAGCACGTGGCGGCTTCTGGGGGCCA
TGTTTGGGGAGGGAGGTGTGCCAGCAGCCTGGAGAGCCTCAGTCCCTGTAGCCCCCTGCCCTGGCACAGCTGCATGC
ACTTCAAGGGCAGCCTTTGGGGGTTGGGGTTTCTGCCACTTCCGGGTCTAGGCCCTGCCCAAATCCAGCCAGTCCTG
CCCCAGCCCACCCACATTGGAGCCCTCCTGCTGCTTTGGTGCCTCAAATAAATACAGATGTCCCC

FIGURE 16

CAGCAGGTCACAGCCCCCTCGAGGCGACAGCGGCCCCGCGCACCAGAGCAGTGGTACAGGCATGGATGGGAAGAAAT
GCAGCGTATGGATGTTCCCTACCTCTTGTATTTACTTTGTTTACTTCAGCTGGATTGTGGATAGTATACTTCATAGCT
GTGGAAGATGACAAAATTTTACCATTAAATTCAGCTGAAAGGAAACCTGGTGTGAAGCATGCACCATATATAAGCAT
TGCAGGTGATGATCCTCCTGCAAGCTGTGTGTTTAGTCAAGTTATGAACATGGCAGCCTTCCTAGCCCTTGTGGTAG
CTGTTCTGCGCTTCATACAACCTGAAACCGAAGTTTTAAACCCGTGGCTGAATATTAGTGGATTGGTGGCTCTGTGT
CTGGCTTCCTTCGGAATGACCTTACTTGGTAATTTTCAGCTCACAAATGATGAAGAAATCCATAACGTCGGAACCTC
CTTGACCTTTGGATTTGGCACATTGACCTGCTGGATCCAGGCTGCGCTGACACTCAAGGTCAACATCAAGAATGAAG
GACGGAGAGTTGGAATTCACGGGTTATTCTGTGCGCATCTATCACTCTCTGTGTGGTCCTCTACTTCATCCTCATG
GCCCCAAGCATCCACATGTATGCAGCCAGGGTCCAGTGGGGCCTGGTCATGTGCTTCCTGTCTATTTTGGCACCTT
TGCCGTGGAGTTCCGGCATTACCGCTATGAGATTGTTTGCTCTGAGTACCAGGAGAATTCCTAAGCTTCTCAGAAA
GCCTGTCAGAAGCTTCTGAATATCAGACTGACCAGGTGTAAACCATCAGTTTTTCCTTGCTGGTGAGGTGGGTGTGA
CAGTGGGGGAGGGGCCAGTAGGACACACTCACAGGACTTGACATAGAACCTCATTTACACACACACACACACACAC
ATTCATGGCCACATTTGCCAAATGAGCTTTTCAGGGCGAGTTATTTCTTTAATGAAAAAGCACAGCCCTTATGTGT
CGAAATACACGCTGTTACACTGAAAATATATGCACGACAGAGCAAGAAGCTTGTGCATGATCACTTCTTATCCGTCC
CCTTCCCAGCACTCCCTCCTCTTCCATTCTCTCCACATGTCTCAAGCACCTACCGAGTAGGGCAGGCCAAATGTT
CCTTGGGAGTAATGCCAACTCCCGACGTTGCCTTCAGGTCCAAAGGGCTTGAACCAGCTCGTGAGGAAGTTCTGAA
TCTGGCACTAATATTCTTGAGTGGATAATAGTGTATCATAGAATAGGACGGAAATTGTATTGAGATGTGACCCTGTG
TCGCCTGTGGAAAGGCATAGTGAGAAGAACTTTCCCACGAAAGCCCCCTTCATCGTTGTTCACTGGTCCGGCTGTGTG
GATCCCAGGAGAGACATATGCCACAGACTGTGAGAGCAAAGCCCGCCGCTGTGATCTGGACTTGATGCACTGTGACT
GAGAATGATTTCCAAATGTGAATATGTGTAGGGACGTGGTCTATCAGGCCTGGAACAAGATGGGGGCAGTGAAGGTA
TGGTTTAGTGTTTGCTTTCATAGTATGCCATGTACAATGTTTTATATTTTCATAGTTTCTTTTAAGTAACTACCATGA
GTCTCTCTAAGCCTCATGGACAAAGATGTAGACCAAATGCAAGAGCTGAGCTTGCTTTGGGTTC AACCATGATCAAA
GAAAAACTGAGGTCACCTGCAGGCTTACGTGGGAAGCTAAGACAATATC

FIGURE 17

CTGCCTCCACTGCTCTGTGCTGGGATCATGGAACTTGCACTGCTGTGTGGGCTGGTGGTGATGGCTGGTGTGATTCC
AATCCAGGGCGGGATCCTGAACCTGAACAAGATGGTCAAGCAAGTGACTGGGAAAATGCCCATCCTCTCCTACTGGC
CCTACGGCTGTCACTGCGGACTAGGTGGCAGAGGCCAACCCAAAGATGCCACGGACTGGTGTGCCAGACCCATGAC
TGCTGCTATGACCACCTGAAGACCCAGGGGTGCGGCATCTACAAGGACTATTACAGATACAACCTTTTCCCAGGGGAA
CATCCACTGCTCTGACAAGGGAAGCTGGTGTGAGCAGCAGCTGTGTGCCTGTGACAAGGAGGTGGCCTTCTGCCTGA
AGCGCAACCTGGACACCTACCAGAAGCGACTGCGTTTCTACTGGCGGCCCCACTGCCGGGGGCAGACCCCTGGGTGC
TAGAAGCCCACACCCTCTACCCTGTTCCCTCAGCATGGAGCTCTGGCATCCCCACCTCAGTATCTAACCTGAACCAGC
CTGGCTTTTCAAACACTCCGGGGGGAGGTAGTCCCAGCCTCCCCCGGAACCCTCTACCAATGCCTTCTGACCTTCTG
AAGCTTTCGGAATCCTCCCAGTTGAGGCAGTAGCTGTGTCTCTGAGGGTGGATGGGAATCTTGGGAGAAGCCCAAG
CAAGGGAGCCCTCAGAGGTGGTGTGTTGGACCAAAGCATCGGGGTGGGGGAGGGGTCTGCCGCTGTCCCCACCTGCT
GGCCCCCTTGTCTTCTCACCCCTCCAATATAGTCTCGGAGCTACAACCTGCAGCAGCCACTATAAAGGGCAATAT
TGATCTTCTGTCCATGTGGCTCTATCTTTTAAACCTCAAGGCCCTCCACTGTCTTAAGATAAAGCCTCTCATAGG
CACTGGGGACCCTGCACAGTCTGGCCATGTGACCCTCTCCCCAGGCAAGCTCTGAAGTCCCTGCAGGTGGAGGCCAT
GCCTGTCTTAAACTCAGTTGCATCCCTGGTGCCCAAAGCAACACCAGAACCAAGAAGGAGCTCCATAAATCCTTCTT
GGGTGAAGCCTAGACAAAGCCGCCAGGTCTTGTGGCTCCAGGCACCAGAGCCTTGAGTACTTTCTCCTGCCTCCAGG
CATTGGCTCAGGGTGAATTACAAGGGGCTACTGAATGGCTATTACTTTCATCACGACTGATCCCCACCTCCTCAGGG
TCAAAGGGCTACTTTCTGGAAGTCTCCCCAGGCTGACTCCTTCTCCCTGACTGCAAGGGCTCACTCCCTCCTCCAAG
CTCCCACAATGCTTCATGGCTCTGCCGCTTACCTAGCTTGGCCTAGAGTGGCAAATGGAACCTTCTCTGATCTCCCC
AACTAGACTGGAGCCCCGAAGGATGGAGACCATGTCTGTGCCATCTCTGTTTCCCCTGTTTTCCACATACTAGGT
GCTCAATTGATGCCTGTGAATGGCGTGAGCCATAATGGATACACAGAGGTTGCAGCAGATGGTGTGGGTACCTCAC
CCAGATATCTTCCAGGCCCAAGGCCCTCTCCCTGAGTGAGGCCAGGTGTTGGCAGCCAACCTGCTCCAATCTGCCTC
CTTCCCCCTAAATACTGCCCTGGTCTAGTGGGAGCTGCCTTCCCCCTGCCCCACCTCTCCCACCAAGAGGCCACCTGT
CACTCATGGCCAGGAGAGTGACACCATGGAGGGTACAATTGCCAGCTCCCCGTGTCTGTGCAGCATTGTCTGGGTT
GAATGACACTCTCAAATTGTTCTCTGGGATCGGGCTGAGGCCAGGCCTCTCCTGGAACCACTCTCTGCTTGGTCTGA
CCCCCTGGCCTATCCAGTTTTCTGTTCCCTCACAGGTTTCTCCAGAAAGTACTCCCTCAGTAAAGCATTTCACACA
AGAAAAAAAAAAAA

FIGURE 18A

GGCTGAAAGAGCCTGAGCTGTGCCTCTCCATTCCACTGCTGTGGCAGGGTCAGAAATCTTGGATAGAGAAAACCTTT
 TGCAAACGGGAATGTATCTTTGTAATTCCTAGCACGAAAGACTCTAACAGGTGTTGCTGTGGCCAGTTACCAACCA
 GCATATCCCCCTCTGCCAAGTGCAACACCCAGCAAAAAATGAAGAGGAAAGCAAACAGGTGGAGACTCAGCCTGAGA
 AATGGTCTGTTGCCAAGCACACCCAGAGCTACCCAACAGATTCTTATGGAGTTCTTGAATTCCAGGGTGGCGGATAT
 TCCAATAAAGCCATGCTATATCCGTGTATCCTATGACACCAAGCCAGACTCACTGCTCCATCTCATGGTGAAAGATTG
 GCAGCTGGAACCCCCAAGCTCTTAATATCTGTGCATGGAGGCCTCCAGAACTTTGAGATGCAGCCCCAAGCTGAAAC
 AAGTCTTTGGGAAAGGCCTGATCAAGGCTGCTATGACCACCGGGGCTGGATCTTCACCGGGGGTGTGAGCACAGGT
 GTTATCAGCCACGTAGGGGATGCCTTGAAAGACCACTCCTCCAAGTCCAGAGGCCGGGTTTGTGCTATAGGAATTGC
 TCCATGGGGCATCGTGGAGAATAAGGAAGACCTGGTTGGAAGGATGTAACAAGAGTGTACCAGACCATGTCCAACC
 CTCTAAGTAAGCTCTCTGTGCTCAACAACCTCCACACCCACTTCTATCCTGGCTGACAATGGCACCCCTGGGCAAGTAT
 GGCGCCGAGGTGAAGCTGCGAAGGCTGCTGGAAAAGCACATCTCCCTCCAGAAGATCAACACAAGACTGGGGCAGGG
 CGTGCCCCCTCGTGGGTCTCGTGGTGGAGGGGGGCCCTAACGTGGTGTCCATCGTCTTGGAATACCTGCAAGAAGAGC
 CTCCCATCCCTGTGGTGATTTGTGATGGCAGCGGACGTGCCTCGGACATCCTGTCTTTGCGCACAAGTACTGTGAA
 GAAGGCGGAATAATAAATGAGTCCCTCAGGGAGCAGCTTCTAGTTACCATTTCAGAAAACATTTAATTATAATAAGGC
 ACAATCACATCAGCTGTTTGAATTATAATGGAGTGCATGAAGAAGAAAGAACTCGTCACTGTGTTTCAGAATGGGT
 CTGAGGGCCAGCAGGACATCGAGATGGCAATTTTAACTGCCCTGCTGAAAGGAACAAACGTATCTGCTCCAGATCAG
 CTGAGCTTGGCACTGGCTTGGAAACCGCGTGGACATAGCACGAAGCCAGATCTTTGTCTTTGGCCCCACTGGACGCC
 CCTGGGAAGCCTGGCACCCCCGACGGACAGCAAAGCCACGGAGAAGGAGAAGAAGCCACCCATGGCCACCACCAAGG
 GAGGAAGAGGAAAAGGGAAAGGCAAGAAGAAAGGGAAAGTGAAAGAGGAAGTGGAGGAAGAACTGACCCCCGGAAG
 ATAGAGCTGCTGAAC TGGGTGAATGCTTTGGAGCAAGCGATGCTAGATGCTTTAGTCTTAGATCGTGTGCACTTTGT
 GAAGCTCCTGATTGAAAACGGAGTGAACATGCAACACTTTCTGACCATTCCGAGGCTGGAGGAGCTCTATAACACAA
 GACTGGGTCCACCAAAACACACTTCATCTGCTGGTGAGGGATGTGAAAAGAGCAACCTTCGGCCTGATTACCACATC
 AGCCTCATAGACATCGGGCTCGTGTGAGGTACCTCATGGGAGGAGCCTACCGCTGCAACTACACTCGGAAAACTT
 TCGGACCCCTTTACAACAACCTGTTTGGACCAAAGAGGCCTAAAGCTCTTAAACTTCTGGGAATGGAAGATGATGAGC
 CTCCAGCTAAAGGGAAGAAAAAAGAAAAAGGAGGAAGAGATCGACATTGATGTGGACGACCCTGCCGTG
 AGTCGGTTCCAGTATCCCTTCCACGAGCTGATGGTGTGGGCAGTGTGATGAAACGCCAGAAAATGGCAGTGTTCTCT
 CTGGCAGCGAGGGGAAGAGAGCATGGCCAAGGCCCTGGTGGCCTGCAAGCTCTACAAGGCCATGGCCCCACGAGTCCT
 CCGAGAGTGATCTGGTGGATGACATCTCCAGGACTTGGATAACAATTCCAAAGACTTCGGCCAGCTTGCTTTGGAG
 TTATTAGACCAGTCTTATAAGCATGACGAGCAGATCGCTATGAAACTCCTGACCTACGAGCTGAAAAACTGGAGCAA
 CTGGACCTGCCTCAAACCTGGCCGTGGCAGCCAAACACCGGGACTTCATTGCTCACACCTGCAGCCAGATGCTGCTGA
 CCGATATGTGGATGGGAAGACTGCGGATGCGGAAGAACCCCGCCCTGAAGGTTATCATGGGGATTCTTCTACCCCCC
 ACCATCTTGTTTTGGAAATTCGCACATATGATGATTTCTCGTATCAAACATCCAAGGAAAACGAGATGGCAAAGA
 AAAAGAAGAGGAAAATACGGATGCAAATGCAGATGCTGGCTCAAGAAAGGGGGATGAGGAGAACGAGCATAAAAAAC
 AGAGAAGTATTTCCCATCGGAACAAAGATCTGTGAATTTCTATAACGCGCCCATTTGTCAAGTTCTGGTTTTACACAATA
 TCATACTTGGGCTACCTGCTGCTGTTTAACTACGTCATCCTGGTGCAGGATGGATGGCTGGCCGTCCCTCCAGGAGTG
 GATCGTCACTCTCTACATCGTGAGCCTGGCGTTAGAGAAGATACGAGAGATCCTCATGTCAGAACCAGGCAAACCTCA
 GCCAGAAAATCAAAGTTTGGCTTCAGGAGTACTGGAACATCACAGATCTCGTGGCCATTTCCACATTTCATGATTGGA
 GCAATTTCTCGCCTACAGAACCAGCCCTACATGGGCTATGGCCGGGTGATCTACTGTGTGGATATCATCTTCTGGTA
 CATCCGTGTCTTGGACATCTTTGGTGTCAACAAGTATCTGGGGCCATACGTGATGATGATTGGAAAGATGATGATCG
 ACATGCTGTACTTTGTGGTTCATCATGCTGGTCTGCTCATGAGTTTTCGGAGTAGCCCGTCAAGCCATTCTGCATCCA
 GAGGAGAAGCCCTCTTGGAACCTGGCCCGAAACATCTTCTACATGCCCTACTGGATGATCTATGGAGAGGTGTTTGC
 AGACCAGATAGACCTCTACGCCATGGAAATTAATCCTCCTTGTGGTGAGAACCTATATGATGAGGAGGGCAAGCGGC
 TTCTCCTGTATCCCCGGCGCCTGGCTCACTCCAGCACTCATGGCGTGCTATCTACTGGTCGCCAACATCCTGCTG
 GTGAACCTGCTGATTGCTGTGTTCAACAATACTTTCTTTGAAGTAAATCAATATCCAACCAGGTGTGGAAGTTCCA
 GCGATATCAGCTGATTATGACATTTTCATGACAGGCCAGTCCCTGCCCCACCGATGATCATTTTAAGCCACATCTACA
 TCATCATTATGCGTCTCAGCGGCCGTGCAGGAAAAAGAGAGAAGGGGACCAAGAGGAACGGGATCGTGGATTGAAG
 CTCTTCTTAGCGACGAGGAGCTAAAGAGGCTGCATGAGTTTCGAGGAGCAGTGCGTGCAGGAGCACTTCCGGGAGAA
 GGAGGATGAGCAGCAGTCTGCCAGCGACGACGCGCATCCGGGTCACTTCTGAAAGAGTTGAAAATATGTCAATGAGGT
 TGGAAGAAATCAATGAAAGAGAACTTTTATGAAAACCTTCCCTGCAGACTGTTGACCTTCGACTTCAGCTAGTGAA
 GAATTATCTAACAGAATGGTGAATGCTTTGAAAACTTGCGGGAATCGACAGGTGTGACCTGATCCAGGCACGGTC
 CCGGGCTTCTTCTGAATGTGAGGCAACGTATCTTCTCCGGCAAAGCAGCATCAATAGCGCTGATGGCTACAGCTTGT
 ATCGATATCATTTTAAACGGAGAAGAGTTATTATTTGAGGATACATCTCTCTCCACGTCACCAGGGACAG

FIGURE 18B

GAGTCAGGAAAAAACCTGTTCCCTTCCGTATAAAGGAAGAGAAGGACGTGAAAAACGCACCTAGTCCCAGAATGTCAG
AACAGTCTTCACCTTTTCACTGGGCACAAGCACATCAGCAACCCAGATGGCAGTCACCTTGCAGTAGATGACTTAAA
GAACGCTGAAGAGTCAAAATTAGGTCCAGATATTGGGATTTCAAAGGAAGATGATGAAAGACAGACAGACTCTAAAA
AAGAAGAACTATTTCCCAAGTTTAAATAAAAACAGATGTGATACATGGACAGGACAAATCAGATGTTCAAAACACT
CAGCTAACAGTGGAAACGACAAATATAGAAGGCACATTTTCCCTATCCCCTGGAAGAAACCAAAATTACACGCTATTT
CCCCGATGAAACGATCAATGCTTGTAACAATGAAGTCCAGAAGCTTCGTCTATTCCCGGGGAAGAAAGCTGGTCG
GTGGGGTTAACCAGGATGTAGAGTACAGTTCATCACGGACCAGCAATTGACGACGGAATGGCAATGCCAAGTTCAA
AAGATCACGCGCTCTCATAGCACAGATATTCCCTTACATTGTGTGCGGAAGCTGCAGTGCAAGCTGAGCAAAAAGAGCA
GTTTGCAGATATGCAAGATGAACACCATGTGCGCTGAAGCAATTCCTCGAATCCCTCGCTTGTCCTAACCATTACTG
ACAGAAATGGGATGGAAACTTACTGTCTGTGAAGCCAGATCAAACCTTTGGGATTCCCATCTCTCAGGTCAAAAAGT
TTACATGGACATCCTAGGAATGTGAAATCCATTCAGGGAAAGTTAGACAGATCTGGACATGCCAGTAGTGTAAGCAG
CTTAGTAATTGTGTCTGGAATGACAGCAGAAAGAAAAAAGGTTAAGAAAGAGAAAGCTTCCACAGAACTGAATGCT
AGTCTGTTTTGTTTTCTTTAATTTTTTTTTTTTAAACAGTCAGAAACCCACTAATGGGTGTCATCTTGGCCCATCCTAAA
CACATMTCCAATTTTCTAAAAACATTTTCCCTTAAAAAATTTTGGAATTCAGACTTGATTTACAATTTAATGCACT
AAAAGTAGTATTTTGTAGXATATGTTAGTAGGCTTAGTTTTTTTCAGTTGCAGTAGTATCAAATGAAAGTGATGATA
CTGTAACGAAGATAAATTGGCTAATCAGTATACAAGATTATACAATCTCTTTATTACTGAGGGCCACCAATAGCCT
AGGAAGTGCCCTCGAGCACTGAAGTCACCATTAGGTCACTCAAGAAGTAAGCAACTAGCTGGGCACAGTGGCTCATG
CCTGTAATCCTAGCACTTTGGGAGGCCAAGGCAGAAAGATAGCTTGAGTCCAGGAGTTTGAGACCAGCCTGGGCAAC
ATAGTGATACCCCATCTCTTAAAAAAAAAAAAAAAAAAAA

FIGURE 19

CTGAATCTTCGTTTCTCTCCCAGGGACCCCTCCATTTTCCATATCCAGGAAAAATGTGATGCGCCACAGGTATCAGCGT
 CTGGATCGCCACTTCACGTTTTAGCCACAAGTGACTCAGTGGAAGATCCAGAGTCAACAGAGGCTCGTCAGGAAGAT
GTCTACAGAAAAGGTAGACCAAAAGGAGGAAGCTGGGGAAAAAGAGGTGTGCGGAGACCAGATCAAAGGACCGGACA
 AAGAGGAGGAACCACCAGCTGCTGCATCCCATGGCCAGGGGTGGCGTCCAGGTGGCAGAGCAGCTAGGAACGCAAGG
 CCTGAACCTGGGGCCAGACACCCTGCTCTCCCGGCCATGGTCAACGACCCTCCAGTACCTGCCTTACTGTGGGCCCA
 GGAGGTGGGCCAAGTCTTGGCAGGCCGTGCCCGCAGGCTGCTGCTGCAGTTTGGGGTGTCTTCTGCACCATCCTCC
 TTTTGCTCTGGGTGTCTGTCTTCCTCTATGGCTCCTTCTACTATTCTATATGCCGACAGTCAGCCACCTCAGCCCT
 GTGCATTTCTACTACAGGACCGACTGTGATTCTCCACCACCTCACTCTGCTCCTTCCCTGTTGCCAATGTCTCGCT
 GACTAAGGGTGGACGTGATCGGGTGTGATGTATGGACAGCCGTATCGTGTTACCTTAGAGCTTGAGCTGCCAGAGT
 CCCCTGTGAATCAAGATTTGGGCATGTTCTTGGTCACCATTTCTTGCTACACCAGAGGTGGCCGAATCATCTCCACT
 TCTTCGCGTTTCGGTGATGCTGCATTACCGCTCAGACCTGCTCCAGATGCTGGACACACTGGTCTTCTCTAGCCTCCT
 GCTATTTGGCTTTGCAGAGCAGAAGCAGCTGCTGGAGGTGGAACCTCTACGCAGACTATAGAGAGAACTCGTACGTGC
 CGACCACTGGAGCGATCATTGAGATCCACAGCAAGCGCATCCAGCTGTATGGAGCCTACCTCCGCATCCACGCGCAC
 TTCCTGGGCTCAGATACCTGCTATACAACCTCCCGATGACCTGCGCCTTCATAGGTGTTGCCAGCAACTTCACCTT
 CCTCAGCGTCATCGTGCTCTTCAGCTACATGCAGTGGGTGTGGGGGGGCATCTGGCCCCGACACCGCTTCTCTTTGC
 AGGTTAACATCCGAAAAAGAGACAATTCCCGGAAGGAAGTCCAACGAAGGATCTCTGCTCATCAGCCAGGGCCTGAA
 GGCCAGGAGGAGTCAACTCCGCAATCAGATGTTACAGAGGATGGTGAGAGCCCTGAAGATCCCTCAGGGACAGAGGG
 TCAGCTGTCCGAGGAGGAGAAACCAGATCAGCAGCCCCTGAGCGGAGAAGAGGAGCTAGAGCCTGAGGCCAGTGATG
 GTTCAGGCTCCTGGGAAGATGCAGCTTTGCTGACGGAGGCCAACCTGCCTGCTCCTGCTCCTGCTTCTGCTTCTGCC
 CCTGTCCTAGAGACTCTGGGCAGCTCTGAACCTGCTGGGGGTGCTCTCCGACAGCGCCCCACCTGCTCTAGTTCCTG
AAGAAAAGGGGCAGACTCCTCACATTCAGCACTTTCCACCTGACTCCTCTCCCCTCGTTTTTCCTTCAATAAACT
 ATTTTGTGTCAGCTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 20A

AACCGGCTGCGGGGATCGGGGCCACCGCTCCCGCTGCTGCTGCTGCTACTGCTGCTGCTGCCGCCACGCGTCCTGCCT
 GCCGCCCCCTTCGTCCGTCCCCCGCGGCCGAGCTCCCGGGGCGTCTGGGCTGCCTGCTCGAGGAGGGCCTCTGCGG
 AGCGTCCGAGGCCGTGTGTGAACGATGGAGTGTGTTGGAAGGTGCCAGAAGGTTCCGGCAATGGACTTTTACCGCTACG
 AGGTGTGCGCCCGTGGCCCTGCAGCGCCTGCGCGTGGCGTTGCAGAAGCTTTCCGGCACAGGTTTCACGTGGCAGGAT
 GACTATACTCAGTATGTGATGGACCAGGAACCTGCAGACCTCCCGAAAACCTACCTGAGGCGTCTGAAGCATCCAG
 CCCAGCCAGGCCCTCAAACACAGCGTTGGCAGCGAGAGGAGGTACAGTCGGGAGGGCGGTGCTGCCCTGGCCAACG
 CCTCCGACGCCACCTGCCCTTCCTGGAGGCCCTGTCCAGGCCCCAGCCTCAGACGTGCTCGCCAGGACCCATACG
 GCGCAGGACAGACCCCCCGCTGAGGGTGATGACCGCTTCTCCGAGAGCATCCTGACCTATGTGGCCACACGTCTGC
 GCTGACGTACCCTCCCGGGCCCCGACCCAGCTCCGGGAGGACCTCTGCCGCGGACCCTCGGCCAGCTCCAGCCAG
 ATGAGCTCAGCCCTAAGGTGGACAGTGGTGTGGACAGACCATCTGATGGCGGCCCTCAGTGCCCTATGCTGCCAG
 AGGCCCCCAGCTCCCCCGGGGAGGGCAGCCTGGAGCCACAGTACCTTCTGCGTGCACCTCAAGAATGCCAGGCC
 TTTGCTGGCACCAGCCGCCCCCAGAAAGTGGCCTTACCTCTGGGAGATTCCGAAGACCCCTCCAGCACAGGCGATG
 GAGCACGGATTCTATACCCTCCTGAAGGACCTGCAGAGGCAGCCGGCTGAGGTGAGGGGCCCTGAGTGGCCTGGAGCTG
 GACGGCATGGCTGAGCTGATGGCTGGCCTGATGCAAGGCGTGGACCATTGGAGTAGCTCGAGGCAGCCCTGGGAGAGC
 GGCGCGGGGCTACATCGTGACAGACAGAGACCCCTGCGCCCCGAGGAAGGAAGGCGGCTGGTGGAGGACGTGCGCC
 TGCAGGACGACGATGATAGACTTTACCAAGAGGTCCACCGTCTGAGTGCCACACTCGGGGGCCTCCTGCAGGACCAC
 GGGTCTCGACTCTTACCTGGAGCCCTCCCCTTTGCAAGGCCCTCGACATGGAGAGGAAGAAGTCCGAGCACCCCTGA
 GTCTTCCCTGTCTTCAAGAGGAGACTGCCGGAGTGGAGAACGTCAAGAGCCAGACGTATTCCAAAGATCTGCTGG
 GGCAGCAGCCGATTCGGAGCCCCGGGGCCGCTGCGTTTGGGGAGCTCCAAAACAGATGCCTGGGCCCTCGAAGGAG
 GAGCAGAGCCTTCCAGCGGGTGTCTAGGAGGCCCTCAGCGACGGCCTGCAATTGGAGGTCCAGCCTTCCGAGGAAGA
 GGCGCGGGGCTACATCGTGACAGACAGAGACCCCTGCGCCCCGAGGAAGGAAGGCGGCTGGTGGAGGACGTGCGCC
 GCCTCCTGCAGGTGCCAGCAGTGCCTTGCCTGACGTGGAGGTTCTCGGACCAGCAGTGACCTTCAAAGTGAGCGCC
 AATGTCCAAAACGTGACCACTGAGGATGTGGAGAAGGCCACAGTTGACAACAAAGACAAACTGGAGGAAACCTCTGG
 ACTGAAAATTCTTCAAACCGGAGTGGGGTGGAAAAGCAAACCTCAAGTTCTTGCCTCCTCAGGCGGAGCAAGAAGACT
 CCACCAAGTTTCATCGCGCTCACCTGGTCTCCCTCGCTGCATCCTGGGCGTCTCCTGGCCTCTGGCCTCATCTAC
 TGCCTCCGCCATAGCTCTCAGCACAGGCTGAAGGAGAAGCTCTCGGGACTAGGGGGCGACCCAGGTGCAGATGCCAC
 TGCCGCTTACCAGGAGCTGTGCCGCCAGCGTATGGCCACGCGGCCACCCAGACCGACCTGAGGGCCCGCACACGTAC
 GCATCAGCAGCGTCTCATCCCAGTTAGCGCAGGGCCGATCCCCAGCCCCCTCCGCACGCAGCAGCGCCTCATCCTGG
 TCCGAGGAGCCTGTGCAGTCCAACATGGACATCTCCACCGGCCACATGATCCTGTCTTACATGGAGGACCACCTGAA
 GAACAAGAACCAGGCTGGAGAAGGAGTGGGAAGCGCTGTGCGCTACCAGGCGGAGCCCCAACAGCTCGTTCTGTGGCCC
 AGAGGGAGGAGAACGTGCCAAGAACCCTCCCTGGCTGTGCTGACCTATGACCACTCCCGGGTCTGTGTAAGGCG
 GAGAAGCCACAGCCACTCAGACTACATCAACGTCAGCCCATATGGATCAGCAGCCGAGGAACCCGCGTACAT
 CGCCACCCAGGGACCGCTGCCCGCCACCGTGGCTGACTTTTGGCAGATGGTGTGGGAGAGCGGCTGCGTGGTGATCG
 TCATGTGACACCCCTCGCGGAGAACGGCGTCCGGCAGTGCTACCACTACTGGCCGGATGAAGGCTCCAATCTCTAC
 CACATCTATGAGGTGAACCTGGTCTCCGAGCACATCTGGTGTGAGGACTTCTGGTGGAGGAGCTTCTATCTGAAGAA
 CCTGCAGACCAACGAGACGCGCACCGTGACGCAGTTCCACTTCTGAGTTGGTATGACCGAGGAGTCCCTTCTCTCT
 CAAGGTCCCTCCTGGACTTCCGCAGAAAAGTAAACAAATGCTACAGGGGGCCGTTCTTGTCCGATAATTGTTTCATTGC
 AGTGACGGTGCAGGCCGGAGCGGCACCTACGTCTGTGATCGACATGGTTCTCAACAAGATGGCCAAAAGGTGCTAAAGA
 GATTGATATCGCAGCGACCCCTGGAGCACTTGGAGGACCAGAGACCCGGCATGGTCCAGACGAAGGAGCAGTTTGAGT
 TCGCGCTGACAGCCGTGGCTGAGGAGGTGAACGCCATCCTCAAGGCCCTTCCCCAGTGAAGCGGCAGCGTCAGGGGCC
 TCAGGGGAGCCCCACCCACGGATGTTGTGAGGAATCATGATCTGACTTTAATTGTGTGCTTCTATTATAACTGC
 ATAGTAATAGGGCCCTTAGCTCTCCCGTAGTCAGCGCAGTTTAGCAGTTAAAAGTGATTTTTGTTTAAATCAAACAA
 TAATAAAGAGAGATTTGTGGAAAAATCCAGTTACGGGTGGAGGGGAATCGGTTTCATCAATTTTCACTTGCTTAAAAA
 AAATACTTTTCTTAAAGCACCCGTTACCTTCTTGGTTGAAGTTGTGTTAACAATGCAGTAGCCAGCACGTTTCGAG
 GCGGTTTCCAGGAAGAGTGTGCTTGTGATCTGCCACTTTCGGGAGGGTGGATCCACTGTGCAGGAGTGGCCGGGGAA
 GCTGGCAGCACTCAGTGAGGCCGCCCCGCACACAAGGCACGTTTGGCATTTCTCTTTGAGAGAGTTTATCATTTGGGA
 GAAGCCGCGGGGACAGAATGAACGTCTGCAGCTTCGGGGCAAGTGAGACAATCACAGCTCCTCGCTGCGTCTCCA
 TCAACACTGCGCCGGGTACCATGGACGGCCCCGTGAGCCACACCGGTGAGCCCAAGCAGAGTGATTAGGGGGCTCCC
 CGGGGGCAGACACCTGTGCACCCCATGAGTAGTGGCCACTTGAGGCTGGCACTCCCTGACCTCACCTTTGCAAGT
 TACAGATGCACCCCAACATTGAGATGTGTTTTTAAATGTTAAATATTGATTTCTACGTTATGAAAACAGATGCCCCC
 GTGAATGCTTACCTGTGAGATAACCACAACCAGGAAGAACAATCTGGGCATTGAGCAAGCTATGAGGGTCCCCGGG
 AGCACACGAACCTGCCAGGCCCGCTGGCTCCTCCAGGCAGTCCCGACCTGTGGGGCCCCAGAGA

FIGURE 20B

GGGGACATTTCCCTCCTGGGAGAGAAGGAGATCAGGGCAACTCGGAGAGGGCTGCGAGCATTTCCCTCCCGGGAGAG
GAAATCAGGGCGACCTGCACGCACTGCGTAGAGCCTGGAAGGGAAGTGAGAAACCAGCCGACCGGCCCTGCCCCTCT
TCCCGGGATCACTTAATGAACCACGTGTTTTGACATCATGTAAACCTAAGCACGTACAGATGATTCCGGATTTGACA
AAATAACATTTGAGTATCCGATTCGCCATCACCCCTACCCCCGAAATAGGACAACCTCACTTCATTGACCAGGATGAT
CACATGGAAGGCGGCGCAGAGGCAGCTGTGTGGGCTGCAGATTTCTGTGTGGGGTTCAGCGTATAAAACGCACCTC
CATCCCGCCCTTCCCACAGCATTCTCTCCATCTTAGATAGATGGTACTCTCCAAAGGCCCTACCAGAGGGAACACGGC
CTACTGAGCGGACAGAATGATGCCAAAATATTGCTTATGTCTCTACATGGTATTGTAATGAATATCTGCTTTAATAT
AGCTATCATTTCTTTTCCAAAATTACTTCTCTTTATCTGGAATTTAATTAATCGAAATGAATTTATCTGAATATAGG
AAGCATATGCCTACTTGTAATTTCTAACTACTTATGTTTGAAGAGAAACCTCCGGTGTGAGATATACAAATATATTT
AATTGTGTCATATTAACTTCCCGGAATTC

FIGURE 21

GCATCTGGTTTGTCTAGATCCGAGAGGCTCTGAAACTGCGGAGCGGCCACCGGACGCCTTCTGGAGCAGGTAGCAGCA
TGCAGCCGCCTCCAAGTCTGTGCGGACGCGCCCTGGTTGCGCTGGTTCTTGCCTGCGGCCTGTGCGGATCTGGGGA
GAGGAGAGAGGCTTCCCGCCTGACAGGGCCACTCCGCTTTTGCAAACCGCAGAGATAATGACGCCACCCACTAAGAC
CTTATGGCCCAAGGGTTCCAACGCCAGTCTGGCGCGGTTCGTTGGCACCTGCGGAGGTGCCTAAAGGAGACAGGACGG
CAGGATCTCCGCCACGCACCATCTCCCCTCCCCCGTGCCAAGGACCCATCGAGATCAAGGAGACTTTCAAATACATC
AACACGGTTGTGTCTGCCTTGTGTTTCGTGCTGGGGATCATCGGGAACCTCCACACTTCTGAGAATTATCTACAAGAA
CAAGTGCATGCGAAACGGTCCCAATATCTTGATCGCCAGCTTGGCTCTGGGAGACCTGCTGCACATCGTCATTGACA
TCCCTATCAATGTCTACAAGCTGCTGGCAGAGGACTGGCCATTTGGAGCTGAGATGTGTAAGCTGGTGCCTTTCATA
CAGAAAGCCTCCGTGGGAATCACTGTGCTGAGTCTATGTGCTCTGAGTATTGACAGATATCGAGCTGTTGCTTCTTG
GAGTAGAATTAAAGGAATTGGGGTTCCAAAATGGACAGCAGTAGAAATTGTTTTGATTTGGGTGGTCTCTGTGGTTTC
TGGCTGTCCCTGAAGCCATAGGTTTTGATATAATTACGATGGACTACAAAGGAAGTTATCTGCGAATCTGCTTGCTT
CATCCCGTTCAGAAGACAGCTTTCATGCAGTTTTTACAAGACAGCAAAAGATTGGTGGCTGTTCAGTTTCTATTTCTG
CTTGCCATTGGCCATCACTGCATTTTTTTTATACACTAATGACCTGTGAAATGTTGAGAAAAGAAAGTGGCATGCAGA
TTGCTTTAAATGATCACCTAAAGCAGAGACGGGAAGTGGCCAAAACCGTCTTTTGCCTGGTCCTTGTCTTTGCCCTC
TGCTGGCTTCCCCTTCACCTCAGCAGGATTCTGAAGCTCACTCTTTATAATCAGAATGATCCCAATAGATGTGAACT
TTTGAGCTTCTGTTGGTATTGGACTATATTGGTATCAACATGGCTTCACTGAATTCCTGCATTAACCCAATTGCTC
TGTATTTGGTGAGCAAAAGATTCAAAAACTGCTTTAAGTCATGCTTATGCTGCTGGTGCCAGTCATTTGAAGAAAAA
CAGTCCTTGAGGAAAAGCAGTCGTGCTTAAAGTTCAAAGCTAATGATCACGGATATGACAACTCCGTTCCAGTAA
TAAATACAGCTCATCTTGAAGAAA

FIGURE 22

MKENYCLQAALVCLGMLCHSHAFAPERRGHLRPSFHGHHEKGKEGQVLQRSKRGWVWNQFFVIEEYTGPDVPLVGRL
HSDIDSGDGNIKYILSGEGAGTIFVIDDKSGNIHATKTLTREERAQYTLMAQAVDRDNRPLEPPSEFIVKVQDIND
NPPEFLHETYHANVPERSNVGTSVIQVTASDADDPYTGNSAKLVYSILEGQPYFSVEAQTGIIRLTALPNMDREAKEE
YHVVIQAKDMGGHMGGLSGTTKVTITLTDVNDNPPKFPQRLYQMSVSEAAVPGEVGRVKAKDPDIGENGLVTYNIV
DGDGMESFEITTDYETQEGVIKLLKPPVDFETERAYSLKVEAANVHIDPKFISNGPFKDTVTVKISVEDADEPPMFLA
PSYIHEVQENAAAGTVVGRVHAKDPDAANSPIRYSIDRHTDLDRFFTINPEDGFIKTTKPLDREETAWLNITVFAAE
IHNRHQEAQVPVAIRVLDVNDNAPKFAAPYEGFICESDQTKPLSNQPIVTISADDKDDTANGPRFIFSLPPEIIHNP
NFTVRDNRDNTAGVYARRGGFSRQKQDLYLLPIVISDGGIPPMSSNTLTITKVCGCDVNGALLSCNAEAYILNAGLS
TGALIAILACIVILLVIVLFLVTLRRQKKEPLIVFEEEDVRENIITYDDEGGGEEDTEAFDIATLQNPDGINGFIPR
KDIKPEYQYMPRPGLRPAPNSVDVDDFINTRIQEADNDPTAPPYDSIQIYGYEGRGSVAGSLSSLESATTDSDLDYD
YLQNWGPRFKKLADLYGSKDTFDDDS

Signal sequence.

amino acids 1-22

Transmembrane domain.

amino acids 617-637

N-glycosylation sites.

amino acids 455-458, 540-543

Glycosaminoglycan attachment sites.

amino acids 83-86, 93-96

N-myristoylation sites.

amino acids 108-113, 215-220, 242-247, 246-251, 247-252, 399-404, 594-599, 599-604, 614-619, 618-623, 749-754, 753-758, 787-792

Cadherins extracellular repeated domain signatures.

amino acids 147-157, 256-266, 476-486

Cadherin cytoplasmic region.

amino acids 641-789

Cadherin domains.

amino acids 59-150, 164-259, 273-375, 388-479, 492-593

FIGURE 23

MYGKSSTRAVLLLLGIQLTALWPAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVTWNFRPLDGGPEQFVF
YYHIDPFQPMSEGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQVKNPPDVGIVIGEIRLSVVHTVRFSEIHFL
ALAIGSACALMIIIVIVVVLFFQHYRKKRWAERAHKVVEIKSKEEERLNQEKKVSVYLEDTD

Signal sequence.

amino acids 1-21

N-glycosylation sites.

amino acids 39-42, 118-121

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 205-208

N-myristoylation sites.

amino acids 15-20, 119-124, 159-164

Immunoglobulin domain.

amino acids 40-125

FIGURE 24

MAEQQGRELEAECPCWNPFNNTFHTPKMLDCCHSFCVECLAHLSLVTPARRRLLCPLCRQPTVLASGQPVTDLPTD
TAMLTLLRLEPHHVILEGHQLCLKDQPKSRYFLRQPRVYTLDLGPQPGGQTGPPPDASATVSTPILIPSHHSLREC
FRNPQFRIFAYLMAVILSVTLLLLIFSIFWTKQFLWGVG

Transmembrane domain.

amino acids 162-182

N-glycosylation site.

amino acids 21-24

N-myristoylation site.

amino acids 68-73

Zinc finger, C3HC4 type, signature.

amino acids 32-41

FIGURE 25

MATAAGATYFQRGSLFWFTVITLSFGYYTWVFWPQSIPYQNLGPLGPFTQYLVDDHHTLLCNGYWLAWLIVGESL
YAIALCKHKGITSGRAQLLWFLQTFFFGIASLTILIA YKRKRQKQT

Transmembrane domain.

amino acids 57-77, 92-112

N-myristoylation site.

amino acids 87-92

Leucine zipper pattern.

amino acids 46-67

FIGURE 26

MASKIGSRRWMLQLIMQLGSVLLTRCPFWGCFSQLMLYAERAEARRKPDIPVPYLYFDMGAAVLCASFMSFGVKRRW
FALGAALQLAISTYAAYIGGYVHYGDWLKVRMYSRTVAIIIGGFLVLASGAGELYRRKPRSRSLQSTGQVFLGIYLIC
VAYSLQHSKEDRLAYLNHLPGGELMIQLFFVLYGILALAFLSGYVTLAAQILAVLLPPVMLLIDGNVAYWHNTRRV
EFWNQMKLLGESVGIFGTAVILATDG

Transmembrane domains.

amino acids 6-26, 51-71, 106-126, 138-158, 174-194, 201-221, 237-256

Glycosaminoglycan attachment site.

amino acids 125-128

N-myristoylation site.

amino acids 245-250

FIGURE 27

GRGSPLALLIRMKTLTLLFGVWALLALILCPGVPEELFEVSIWPSQALVEFGQSLVCNCSTTCPDGPGSGIETFLKKTO
VDKGPQWKEFLLEDVTENSILQCFFSCAGIQKDTSLGITVYQPPEQVILELQPAWVAVDEAFTVKCHVPSVAPLES
TLALLQGNQELHRKNFTSLAVASQRAEVIISVRAQKENDRCNSSCHAE DL SLQGGRLFGQSSPIRIVRIFEFSQSP
HIWVSSLLLEAGMAETVSCEVARVFPAKEVMFHFLEDQELSSSLSWEGDTAWANATIRTMEAGDQELSCFASLGAME
QKTRKLVHSYSFPPPILELKESYPLAGTDINVTCSGHVLTSPSPTLRLQGAPDLPAGEPAWLLLTAREEDDGXNFSC
EASLVVQGQRLMKT TVIQLHILKPQLEESSCPGKQ TWLEGMEHTLACVPGKNPAPALVCTWNGVVFDLEVPQKAT

Signal sequence.

amino acids 1-30

N-glycosylation sites.

amino acids 56-59, 169-172, 196-199, 285-288, 339-342, 382-385

N-myristoylation sites.

amino acids 3-8, 242-247, 335-340, 380-385, 425-430

Intercellular adhesion molecule (ICAM) homology.

amino acids 17-123

FIGURE 28

MLPRLLLLICAPLCEPAELFLIASPSHPTGSPVTLTCKMPFLQSSDAQFQFCFFRDTRALGPGWSSSPKLQIAAMW
KEDTGSYWCEAQTMASKVLRSSRSQINVHIPVSRPILMLRAPRAQAAVEDVLELHCEALRGSPPILYWIFYHEDITLG
SRSAPSGGGASFNLSTEEHSGNYSCEANGLGAQRSEAVTLNFTVPTGARSNHLTSGVIEGLLSTLGPATVALLFC
YGLKRKIGRRSARDPLRSLPALPQEFTYLNSTPGQLQPIYENVNVVSGDEVYSLAYYNQPEQESVAAETLGTHMED
KVSLDIYSRLRKANITDVDYEDAM

Signal sequence.

amino acids 1-16

Transmembrane domain.

amino acids 211-231

N-glycosylation sites.

amino acids 167-170, 177-180, 197-200, 322-325

Glycosaminoglycan attachment site.

amino acids 160-163

N-myristoylation sites.

amino acids 31-36, 82-87, 161-166, 163-168, 176-181, 187-192, 203-208, 212-217,
216-221

Amidation site.

amino acids 238-241

Immunoglobulin domains.

amino acids 31-88, 126-182

FIGURE 29

MLLWASLLAFAPVCGQSAAAHKPVISVHPPWTTFFKGERVTLTCNGFQFYATEKTTWYHRHYWGEKLTLTTPGNTLEV
 RESGLYRCQARGSPRSNPVRLLFSSDSLILQAPYSVFEGDTLVLRCHRRRKEKLTAVKYTWNGNILSISNKSWDLLI
 PQASSNNNGNYRCIGYGDENDVFRSNFKI I KIQELFPHPELKATDSQPTEGNSVNLSCETQLPPERSDTPLHFNFFR
 DGEVILSDWSTYPELQLPTVWRENSGSYWCGAETVRGNIHKHSPSLQIHVQRI PVSGVLLETQPSGGQAVEGEMLV
 VCSVAEGTGDTTFSWHREDMQESLGRKTQRSRLRAELELPAIRQSHAGGYCTADNSYGPVQSMVLNVTVRETPGNRD
 GLVAAGATGGLLSALLLAVALLFHCWRRRKSGVGFLGDETRLPPAPGPGESSHSICPAQVELQSLYVDVHPKKGDLV
 YSEIQTTQLGEEEEANTSRTLLEDKDVSVVYSEVKTQHPDNSAGKISSKDEES

Signal sequence.

amino acids 1-15

Transmembrane domain.

amino acids 387-407

N-glycosylation sites.

amino acids 147-150, 209-212, 374-377, 478-481

Glycosaminoglycan attachment site.

amino acids 416-419

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 413-416

Tyrosine kinase phosphorylation site.

amino acids 457-463

N-myristoylation sites.

amino acids 15-20, 81-86, 89-94, 140-145, 163-168, 205-210, 257-262, 315-320,
 355-360, 382-387, 386-391, 391-396, 394-399, 395-400

Amidation site.

amino acids 332-335

Immunoglobulin domains.

amino acids 37-87, 116-169, 205-263, 303-361

FIGURE 30

MTVIRFFPAASATKRVLPVLRVSSPRTWNPVPESPRIAPAPRLPKRMSGAPTAGAALMLCAATAVLLSAQGGPVQS
KSPRFASWDEMNVLAHGLLQLGQGLREHAERTRSQLSALERRLSACGSACQGTEGSTDLPLAPESRVDPEVLHSLQT
QLKAQNSRIQQLFHKVAQQQRHLEKQHLRIQHLSQSFGLLDHKHLDHEVAKPARRKRLPEMAQPVDPAHNVSRLHRL
PRDCQELFQVGERQSGLFEIQPGSPPFLVNCKMTSDGGWTVIQRRHDGSVDFNRPWEAYKAGFGDPHGEFWLGLEK
VHSITGDRNSRLAVQLRDWDGNAELLQFSVHLGGEDTAYSLQLTAPVAGQLGATTVPVPSGLSVPFSTWDQDHLRRD
KNCAKSLSGGWVFGTCSHSNLNGQYFRSIPQQRQKLKKGIFWKTRGRYYPLQATTMLIQPMAAEAAS

Transmembrane domain.

amino acids 49-69

N-glycosylation site.

amino acids 224-227

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 46-49, 118-121

N-myristoylation sites.

amino acids 50-55, 129-134, 341-346, 357-362

Fibrinogen beta and gamma chains C-terminal domain signature.

amino acids 396-408

Fibrinogen beta and gamma chains.

amino acids 231-447

FIGURE 31

MLLWVILLVLAPVSGQFARTPRPIIFLQPPWTTVFQGERVTLTCKGFRFYSPQKTKWYHRYLGKEILRETPDNILEV
QESGEYRCQAQGSPLSSPVHLDFSSEMGFPHAAQANVELLGSSDLLT

Signal sequence.

amino acids 1-15

N-myristoylation site.

amino acids 89-94

FIGURE 32

MAGPRPSPWARLLAALISVSLSGTLANRCKKAPVKSCTECVRVKDKCAYCTDEMFRDRRCNTQAELLAAGCQRESI
VVMESFFQITEETQIDTTLRRSQMSPQGLRVRLRPGEERHFELEVFEPLSPVDLYILMDFSNSMSDDLDNLKKMGQ
NLARVLSQLTSDYTIIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPPFSFKNVISLTEDVDEFRNKLQGERISGNLDA
PEGGFDAILQTAVCTRDIGWRPDSTHLLVFSTESAFHYEADGANVLGIMSRNDERCHLDTTGTYTQYRTQDYPSVP
TLVRLLAKHNIIPFAVTNYSYSYIEKLHTYFPVSSSLGVLQEDSSNIVELLEAFNRIRSNLDIRALDSPRGLRTEV
TSKMFQKTRTGSFHIRRGVEGIYQVQLRALEHVDGTHVCQLPEDQKGNHILKPSFSDGLKMDAGIICDVCTCELQKE
VRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDIQPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYD
NFQCPRTSGLCNDRGRCSMGQCVCPEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGRCHCHQQSLYTDITICEI
NYSIAIHPLCEDLRSCVQCQAWGTGEKKGRTEECNFKVKMVDLKRAEVVRCSFRDEDDDDCTYSYTMEGDGAPG
PNSTVLVHKKKDCPPGSFWWLIPLLLLLLPLALLLLLCWKYCACCKACLALLPCCNRGHMVGFKEDHYMLRENLMA
SDHLDTPLMRSGNLKGRDVRWKVTNNMQRPGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRCAQLRQEV
EENLNEVYRQISGVHKLQQTFRQQPNAGKKQDHTIVDTVLMAPRSAPKALLKLTQVEQRAFHDLKVAPGYTTLT
ADQDARGMVEFQEGVELVDVRVPLFIRPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFSV
SRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAGQNRDYIPVEGELLFQPGAEWKEQVKLLELQEVDSLLRGRQVRR
FHVQLSNPKFGAHLGQPHSTTIIIRDPELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNLWLPSPGKPM
GYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYGAQGEQPYSSSLVSCRTHQVEVPSEPGRLAFNVVS
STVTQLSWAEPATNGEITAYEVCYGLVNDNRPPIGPMKKVLVDNPNKRMLLIENLRESQPYRYTVKARNGAGWGPE
REAIINLATQPKRPMSPPIIPDIPVDAQSGEDYDSFLMYSDDLRSPSGSQRPVSDDTEHLVNGRMDFAFPGSTN
SLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNLSLTRSEHSHSTTLPRDYSTLTSVSSHDSRLTAGVPDTPTR
LVFSALGPTSLRVSWQEPRCERPLQGYSEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYVFRVRAQSQEGWGR
EREGVITIESQVHPQSPCLPLPGSAFTLSTPSAPGPLVFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPA
TAFRVDGDSPE SRLTPGLSENVYPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGSRAGLFQHPLOQSEYSSIT
TTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVTQEFVSRTLTSGLTSTHMDQQFFQT

Signal sequence.

amino acids 1-23

Transmembrane domain.

amino acids 711-731

N-glycosylation sites.

amino acids 327-330, 491-494, 579-582, 617-620, 695-698, 980-983, 1523-1526

Glycosaminoglycan attachment site.

amino acids 513-516

Tyrosine kinase phosphorylation sites.

amino acids 43-50, 401-408, 674-682, 1596-1604

N-myristoylation sites.

amino acids 24-29, 153-158, 234-239, 420-425, 477-482, 487-492, 560-565,
587-592, 588-593, 782-787, 1031-1036, 1089-1094, 1093-1098, 1383-1388,
1563-1568, 1613-1618, 1614-1619, 1635-1640, 1676-1681, 1719-1724, 1739-1744

Amidation sites.

amino acids 875-878, 974-977

Cell attachment sequence.

amino acids 1003-1006

EGF-like domain cysteine pattern signatures.

amino acids 479-490, 562-573

Integrins beta chain cysteine-rich domain signature.

amino acids 512-525, 590-603

Integrin, beta chain homology.

amino acids 37-455

Calx-beta domain.

amino acids 979-1084

Fibronectin type III domain.

amino acids 1127-1208, 1220-1310, 1458-1542, 1571-1658

FIGURE 33

MTLDRPGEGATMLKTFTVLLFCIRMSLGMTSIVMDPQPELWIESNYPQAPWENITLWCRSPSRISSEKFLLLKDKTQM
 TWIRPSHKTFFQVSFLIGALTESNAGLYRCCYWKETGWSKPSKVLELEAPGQLPKPIFWIQAETPALPGCNVNILCHG
 WLQDLVFMFLKEGYAEPVDYQVPTGTMAIFSIDNLTPEDEGVYICRTHIQMLPTLWSEPSNPLKLVVAGLYPKPTLT
 AHPGPIMAPGESLNLRCQGPIYGMTFALMRVEDLEKSFYHKKTIKNEANFFFQSLKIQDTGHYLCFYDASYRGSLL
 SDVLKIWVTDTFPKTWLLARPSAVVQMGQNVSLRCRGPVDGVGLALYKKGEDKPLQFLDATSIDDNTSFFLNNTYS
 DTGIYSCHYLLTWKTSIRMPSHNTVELMVVDKPPKPSLSAWPSTVFKLGKAITLQCRVSHPVLEFSLEWEERETFQR
 FSVNGDFIISNVDGKGTGTYSYRVETHPNMWSHRSEPLKLMGPAGYLTWNYVLNEAIRLSLIMQLVALLLVVLWI
 RWKCRRLRIREAWLLGTAQGVMTLFIIVTALLCCGLCNGVLIETEIVMPTPKPELWAETNFPAPWKNLTLWCRSPS
 GSTKEFVLLKDGTDGWIATRPASEQVRAAFPLGALTQSHTGSYHCHSWEEMAVSEPSEALELVGTDILPKPVISASPT
 IRGQELQLRCKGWLAGMGFALYKEGEQEPVQQLGAVGREAFFTIQRMEDKDEGNYSRTHTEKLPFKWSEPSEPLEL
 VIKEMYPKPFFKTWASPVVTPGARVTFNCSTPHQHMSFILYKDGSEIASSDRSWASPGASAAHFLIISVGIGDGGNY
 SCRYYDFSIWSESPDPVELVVTEFYPKPTLLAQPGPVVFPKGSVILRCQGTFFQGMRFALLQEGAHVPLQFRSVSGNS
 ADPLLHTVGAEDSGNYSYIYETTMNSNRGSYLSMPLMIWVTDTFPKPWLFAEPSSVPMGQNVTLWCRGPVHGVGI
 LHKEGEATSMQLWGSTSDNGAFPITNISGTSMGYSYCCYHPDWTSSIKIQPSNTLELLVTGLLPKPSLLAQPGPMVA
 PGENMTLQCQGLPDSTFVLLKEGAQEPLEQQRPSPGYRADFWMPAVRGEDSGIYSCVYLDSTPFAASNHSLSLEIW
 VTDKPPKPSLSAWPSTMFKLKGDITLQCRGPLPGVEFVLEHDGEEAPQQFSEDGDFVINNVEGKGIGNYSCSYRLQA
 YPDIWSESPDPLELVGAAGPVAQECTVGNIVRSSLIIVVVVALGVVLAIEWKKWPRRLRTRGSETDGRDQTIALEECN
 QEGEPGTPANSPSSTSQRISVELPVPI

Signal sequence.

amino acids 1-28

Transmembrane domains.

amino acids 517-537, 555-575, 1261-1281

N-glycosylation sites.

amino acids 53-56, 338-341, 374-377, 381-384, 607-610, 747-750, 798-801,
 846-849, 939-942, 986-989, 1027-1030, 1082-1085, 1147-1150, 1223-1226

Tyrosine kinase phosphorylation sites.

amino acids 287-295, 1125-1132

N-myristoylation sites.

amino acids 102-107, 145-150, 195-200, 254-259, 305-310, 336-341, 349-354,
 388-393, 480-485, 555-560, 573-578, 656-661, 709-714, 746-751, 792-797, 814-819,
 828-833, 840-845, 844-849, 845-850, 897-902, 901-906, 933-938, 938-943, 953-958,
 984-989, 1015-1020, 1030-1035, 1130-1135, 1222-1227, 1276-1281, 1315-1320

Immunoglobulin domains.

amino acids 51-108, 145-201, 241-298, 336-394, 434-486, 605-662, 696-752,
 792-851, 888-944, 984-1040, 1080-1136, 1176-1228

FIGURE 34

MAPEPAPGRMTMVPLVPALVMLGLVAGAHGDSKPVFIKVPEDQTGLSGGVASFVCQATGEPKPRITWMKKGKKVSSQR
 FEVIEFDDGAGSVLRIQPLRVQRDEAIYECTATNSLGEINTSAKLSVLEEEQLPPGFPSIDMGPQLKVVEKARTATM
 LCAAGGNPDPEISWFKDFLPVDPATSNRIKQLRSGALQIESSEESDQGYECVATNSAGTRYSAANLYVRVRRVA
 PRFSIPPSSQEVMPGGSVNLTCVAVGAPMPYVKWMMGAEELTKEDEMPVGRNVLELSNVVRSANYTCVAISSLGMIE
 ATAQVTVKALPKPPIDLVVTETTATSVTLTWDSGNSEPVTTYGIQYRAAGTEGPFQEVVDGVATTRYSIGGLSPFSEY
 AFRVLAVNSIGRGPPEAVRARTGEQAPSSPPRRVQARMLSASTMLVQWEPPEEPNGLVRGYRVYYTPDSRRPPNAW
 HKHNTDAGLLTTVGSLPLGITYSLRVLAFTAVGDGPPSPITQVKTQQGVPAQPADFQAEVESDTRIQLSWLLPPQER
 IIMYELVYWAAEDEDQQHKVTFDPTSSYTTLEDLKPDTLYRFQLAARSDMGVGVFTPTIEARTAQSTPSAPPQKVMCV
 SMGSTTVRVSWVPPPADSRNGVITQYSVAHEAVDGEDRGRHVVDGISREHSSWDLVGLEKWTEYRVWVRAHTDVGPG
 PESSPVLVRTDEDVPSGPPRKVEVEPLNSTAVHVYWKLPVPSKQHGQIRGYQVTYVRLLENGEPRGLPIIQDVMLAEA
 QWRPEESEDYETTISGLTPETTSVTVAAAYTTKGDGARSKPKIVTTTGAVPGRPTMMISTTAMNTALLQWHPPKELP
 GELLGYRLQYCRADEARPNTIDFGKDDQHFTVTGLHKGTYYIFRLAAKNRAGLGEEFEKEIRTPEDLPSGFPQNLHV
 TGLTTSTTELAWDPPVLAERNRIISYTVVFRDINSQOELQNITTDTRFTLTGLKPDTTYDIKVRATWSKSGSPLSP
 SIQSRTPVEQVFAKNFRVAAAMKTSVLLSWEVPDSYKSAVPFKILYNGQSVEVDGHSRKLIALDLPNTEYSFVLM
 NRGSSAGLQHLVSIRTAPDLLPHKPLPASAYIEDGRFDLSMPHVQDPSLVRWFYIVVVPIDRVGGSMLTPRWSTPE
 ELELDELLEAIEQGGEEQRRRRRQAERLKPYPVAAQLDVLPEFTTLGDKKNYRGFYNRPLSPDLSYQCFVLASLKEPM
 DQKRYASSPYSDEIVVQVTPAQQQEPEMLWVTGPVLAVILILIVIAILLFKRKRTHSPSSKDEQSIGLKDSLLAH
 SSDPVEMRRLNYQTGMRDHPPIIPITDLADNIERLKANDGLKFSQEYESIDPGQQFTWENSNLEVNKPKNRYANVIA
 YDHSRVILTSIDGVPGSDYINANYIDGYRKQAYIATQGGLPETMGDFWRMVWEQRTATVVMTRLEEKSRVKCDQY
 WPARGTETCGLIQVTLTDLTVELATYTVRTFALHKSGSSEKRELRFQFMAWPDHGVPEYPTPILAFLLRRVKACNPLD
 AGPMVVHCSAGVGRGTCFIVIDAMLERMKHEKTVDIYGHVTCMRSQRNMYVQTEDQYVFIHEALLEAATCGHTEVPA
 RNLYAHIQKLGQVPPGESVTAMELEFKLLASSKAHTSRFISANLPCNKFKNRLVNMIPYELTRVCLQPIRGVEGSDY
 INASFLDGYRQQKAYIATQGGLAESTEDFWRMLWEHNSTIIVMLTKLREMGREKCHQYWPASERSARYQYFVVDPMAE
 YNMPQYILREFKVTDARDGQSRTIRQFQFTDWPEQGVPKTGEGFIDFIGQVHKTKEQFGQDGPITVHCSAGVGRGTGV
 FITLSIVLERMRYEGVDMFQTVKTLRTQRPAMVQTEDQYQLCYRAALEYLGSFDHYAT

Signal sequence.

amino acids 1-29

Transmembrane domain.

amino acids 1262-1282

N-glycosylation sites.

amino acids 117-120, 250-253, 295-298, 721-724, 966-969, 1696-1699, 1731-1734

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 71-74, 1286-1289

Tyrosine kinase phosphorylation sites.

amino acids 97-105, 773-780, 1229-1237, 1687-1694

N-myristoylation sites.

amino acids 22-27, 44-49, 47-52, 214-219, 246-251, 305-310, 368-373, 442-447,
 470-475, 481-486, 739-744, 818-823, 881-886, 926-931, 1081-1084, 1301-1306,
 1349-1354, 1362-1367, 1399-1404, 1468-1473, 1551-1556, 1688-1693, 1790-1795,
 1807-1812, 1842-1847, 1847-1852

Amidation site.

amino acids 69-72

Leucine zipper pattern.

amino acids 1262-1283

Myb DNA-binding domain repeat signature 1.

amino acids 1151-1159

Tyrosine specific protein phosphatases active sites.

amino acids 1546-1558, 1837-1849

Immunoglobulin domains.

amino acids 47-109, 149-209, 246-300

Fibronectin type III domains.

amino acids 319-401, 413-500, 512-594, 606-696, 708-809, 821-904, 915-1000

Protein-tyrosine phosphatase homology.

amino acids 1375-1606, 1664-1897

FIGURE 35

MRRAALWLWLCALALSLQLALPQIVATNLPPEDQDGSDDSDNFSGSGAGALQDITLSQQTPSTWKDTQLLTAIPTS
 PEPTGLEATAASTSTLPAGEGPKGEAVVLPEVEPGLTAREQEATPRPRETTQLPTTHQASTTTATTAQEPATSHPH
 RDMQPGHHETSTPAGPSQADLHTPHTEDGGPSATERAAEDGASSQLPAAEGSGEQDFTFETSGENTAVVAVEPDRRN
 QSPVDQGATGASQGLLDRKEVLGGVIAGGLVGLIFAVCLVGFMLYRMKKKDEGSYSLEEPKQANGGAYQKPTKQEEFYA

Signal sequence.

amino acids 1-22

Transmembrane domain.

amino acids 252-272

N-glycosylation site.

amino acids 43-46

Glycosaminoglycan attachment sites.

amino acids 45-48, 47-50

Tyrosine kinase phosphorylation site.

amino acids 279-286

N-myristoylation sites.

amino acids 46-51, 82-87, 183-188, 238-243, 241-246, 254-259, 255-260, 259-264,
 263-268

Syndecans signature.

amino acids 276-289

Syndecan domain.

amino acids 3-308

FIGURE 36

MKIFLPVLLAALLGVERASSLMCFSCLNQKSNLYCLKPTICSDQDNYCVTVSASAGIGNLVTFGHSLSKTCSPACPI
PEGVNVGVASMGISCCQSFLCNFSAADGGLRASVTLLGAGLLLSLLPALLRFGP

Signal sequence.

amino acids 1-20

Transmembrane domain.

amino acids 108-128

N-glycosylation site.

amino acids 99-102

N-myristoylation sites.

amino acids 14-19, 58-63, 80-85, 89-94, 105-110, 106-111, 117-122

u-PAR/Ly-6 domain.

amino acids 21-100

FIGURE 37

MDGKKCSVWMFLPLVFTLFTSAGLWIVYFIAVEDDKILPLNSAERKPGVKHAPYISIAGDDPPASCVFSQVMNMAAF
LALVVAVLRFIQLKPKVLNPWLNISGLVALCLASFGMTLLGNFQLTNDEEIHNVGTSLTFGFGTLTCWIIQAALTLKV
NIKNEGRRVGIPRVILSASITLCVVLYFILMAQSIHMYAARVQWGLVMCFLSYFGTFAVEFRHYRYEIVCSEYQENF
LSFSESLSEASEYQTDQV

Transmembrane domains.

amino acids 7-27, 65-85, 97-117, 131-151, 166-186

N-glycosylation site.

amino acids 100-103

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 4-7

N-myristoylation sites.

amino acids 132-137, 140-145, 199-201

Amidation sites.

amino acids 2-5, 159-162

FIGURE 38

MELALLCGLVVMAGVPIPIQGGILNLNKMVKQVTGKMPILSYWPYGCHCGLGGRGQPKDATDWCCQTHDCCYDHLKTQ
GCGIYKDYYRYNFSQGNIHCSDKGSWCEQQLCACDKEVAFCLKRNLDTYQKRLRFYWRPHCRGQTPGC

Signal sequence.

amino acids 1-17

N-glycosylation site.

amino acids 89-92

N-myristoylation sites.

amino acids 20-25, 45-50, 93-98, 140-145

Phospholipase A2 histidine active site.

amino acids 63-70

Phospholipase A2 aspartic acid active site.

amino acids 108-118

Phospholipase A2 homology.

amino acids 21-145

FIGURE 39

MYIRVSYDTKPDSSLHLMVKDWQLELPKLLISVHGGLQNFEMQPKLKQVFGKGLIKAAMTTGAWIFTGGVSTGVISH
 VGDALKDHSKSRGRVCAIGIAPWGIVENKEDLVGKDVTRVYQTMSNPLSKLSVLNNSHTHFIADNGTLGKYGAEV
 KLRRLLEKHISLQKINTRLGQGVPLVGLVVEGGPNVVSIVLEYLQEEPPIPVVICDGSGRASDILSFAHKYCEEGLI
 INESLREQLLVTIQKTFNYNKAQSHQLFAIMECMKKKELVTVFRMGSEGQODIEMAILTALLKGTNVSAPDQLSLA
 LAWNRVDIARSQIFVFGPHWTPGLSLAPPTDSKATEKEKKPPMATTKGGRGKGKGGKGGKGVKEVEEETDPRKIELL
 NWNVALEQAMLDALVLDVRDVFVKLLIENGVMQHFLLTIPRLEELYNTRLGPPNTHLLVLDVKKSNLPPDYHISLID
 IGLVLEYLMGGAYRCNYTRKNFRTLYNNLFGPKRPKALKLLGMEDDEPPAKGKKKKKKKKKEEIDIDVDDPAVSRFQ
 YPFHELMVWAVLMKRQKMAVFLWQRGEESMAKALVACKLYKAMAHESSDLDVDDISQDLNNSKDFGQLALELLDQ
 SYKHDEQIAMKLLTYELKNWSNSTCLKLAVAACHRDFIAHTCSQMLLTDMMWGRRLMRKNPGLKVIMGILLPPTILF
 LEFRTYDDFSYQTSKENEDGKEKEEENTDANADAGSRKGDEENEHKKQRSIPIGTKICEFYNAPIVKFWFYTISYLG
 YLLLFNYVILVRMDGWPSLQEWIVISYIVSLALEKIREILMSEPGKLSQKIKVWLQEYWNITDLVAISTFMIGAILR
 LQNQPYMGYGRVIYCVDIIFWYIRVLDIFGVNKYLGPPYMMIGKMMIDMLYFVVMILVVLMSFGVARQAILHPEEK
 SWKLARNIFYMPYWMIYGEVFADQIDLYAMEINPPCGENLYDEEGKRLPPCIPGAWLTPALMACYLLVANILLVNL
 IAVFNNTFFEVKSISNQVWKFQRYQLIMTFHDRPVLPPPMIILSHIYIIIMRLSGRCRKKREGDQEERDRGLKFLS
 DEELKRLHEFEEQCQVQEHFREKEDEQQSSSDERIRVTSERVENMSMRLEEINERETFMKTSQTVDLRLAQLEELSN
 RMVNALENLAGIDRSDLIQARSRASSECEATYLLRQSSINSADGYSLYRYHFNGEELLFEDTSLSTSPGTGVRKKTC
 SFRIKEEKDVKTHLVPECQNSLHLSLGTSTSATPDGSHLAVDDLKNAEESKLGPDIGISKEDDERQTDKKEETISP
 SLNKTDVIHGQDKSDVQNTQLTVETTNIEGTISYPLEETKITRYFPDETINACKTMKSRSFVYSRGRKLVGQVNDV
 EYSSITDQQLTTEWQCQVQKITRSHSTDIPIYVSEAAVQAEQKEQFADMQDEHHVAEAIPIRRLSLTITDRNGMEN
 LLSVKPDQTLGFPSLSRSLHGHPRNVKSIQGKLDKDRSGHASSVSSLVIVSGMTAEKKVKKEKASTETEC

Transmembrane domains.

amino acids 759-779, 828-848, 857-877, 893-913, 976-996, 992-1012, 1031-1051

N-glycosylation sites.

amino acids 133-136, 144-147, 233-236, 298-301, 478-481, 601-604, 635-638,
 638-641, 830-833, 1006-1009, 1121-1124, 1312-1315

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 1228-1231

Tyrosine kinase phosphorylation sites.

amino acids 697-704, 891-898

N-myristoylation sites.

amino acids 35-40, 53-58, 68-73, 69-74, 102-107, 211-216, 229-234, 296-301,
 473-478, 728-733, 747-752, 1166-1171, 1259-1264, 1268-1273, 1319-1324

Amidation sites.

amino acids 362-365, 513-516, 968-971, 1374-1377

Ion transport protein homology.

amino acids 789-1005

FIGURE 40

MSTEKVDQKEEAGEKEVCGDQIKGPDKEEEPAAASHGQGWRPGGRAARNARPEPGARHPALPAMVNDPPVPALLWA
 QEVGQVLAGRARRLLQLQFGVLFCTILLLLWVSFVFLYGSFYYSYMPVTVSHLSPVHFYYRTDCDSSTTSLCSFPVANVS
 LTKGGRDRVLMYGQPYRVTLLELPESPVNQDLGMFLVTISCYTRGGRIISTSSRSVMLHYRSDLLQMLDTLVFSSL
 LLFGFAEQQLLEVELYADYRENSYVPTTGAIIEIHSKRIQLYGAYLRIHAHFTGLRYLLYNFPMTCAFIGVASNFT
 FLSVIVLFSYMQVWGGIWPRHRFSLQVNIRKRDNSRKEVQRRISAHQPGPEGQEESTPQSDVTEDEGESPEDPSGTE
 GQLSEEEKPDQQPLSGEEELEPEASDGSGSWEDAALLTEANLPAPAPASASAPVLETLGSSSEPAGGALRQRPTCSSS

Transmembrane domains.

amino acids 94-114, 134-154, 292-312, 308-328

N-glycosylation sites.

amino acids 152-155, 306-309

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 350-353

Tyrosine kinase phosphorylation site.

amino acids 240-248

N-myristoylation sites.

amino acids 44-49, 81-86, 96-101, 302-307, 361-366

Leucine zipper patterns.

amino acids 84-105, 91-112

FIGURE 41

MGPPPLLLLLLLLLLLLLPPRVLPAPSSVPRGRQLPGR LGCLLEGLCGASEACVNDGVFGRCQKVPAMDFYRYEVS PV
 ALQRLRVALQKLSGTGFTWQDDYTQYVMDQELADLPKTYLRRPEASSPARPSKHSVGSERRYSREGGAALANALRRH
 LPFLEALSQAPASDVLARTHTAQDRPPAEGDDRFSESILTYVAHTSALTYPGPRTQLREDLLPRTLGLQLQPDLS P
 KVDSGVDRHHLMAALSAYAAQRPPAPPGEGLSLEPQYLLRAPSRMPRPLLAPAAPQKWPSPLGDS EDPSSSTGDGARIH
 TLLKDLQRQPAEVRGLSGLELDGMAELMAGLMQGV D HGVARGSPGRAALGESGEQADGPKATLRGDSFPDDGVQDDD
 DRLYQEVHRLSATLGGLLDHGSRLLP GALPFARPLDMERKKSEHPSSLSSEEETAGVENVKSQTYSKDLLGQQPH
 SEPGAAAFGELQNQMPGPSKEEQSLPAGAEALSDGLQLEVQPSEEEARGYIVTDRDPLRPEEGRRRLVEDVARLLQV
 PSSAFADVEVLGPAVTFKVSANVQNVTTEDVEKATVDNKKLEETSGLKILQTGVGSKSKLFLPPQAEQEDSTKFI
 ALTLVSLACILGVLLASGLIYCLRHSSQHR LKEKLSGLGGDPGADATAAYQELCRQRMATRPDRPEGPHTSRISSV
 SSQFSDGPIPSPSARSSASSWSEEPVQSNMDISTGHMILSYMEDHLKKNRLEKEWEALCAYQAEPNSSFVAQREEN
 VPKNRSLAVLTYDHSRVLLKAENSHSHSDYINAS PIMDHDPNPAYIATQGPLPATVADFWQMVWESGCVVIVMLTP
 LAENGVRQCYHYWPDEGSNLYHIYEVNLVSEHIWCEDFLVRSFYLNKLNQTNETRTVTQFHFLSWYDRGVPSSSRLL
 DFRRKVNKC YRGRSCPIIVHCSDGAGRSGTYVLIDMVLNKMAGKAKEIDIAATLEHLRDQRPGMVQTKEQFEFALTA
 VAEEVNAILKALPQ

Signal sequence.

amino acids 1-21

Transmembrane domain.

amino acids 616-636

N-glycosylation sites.

amino acids 564-567, 760-763, 774-777, 898-901

Glycosaminoglycan attachment sites.

amino acids 90-93, 652-655

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 137-140, 425-428

N-myristoylation sites.

amino acids 44-49, 47-52, 91-96, 144-149, 338-343, 342-347, 346-351, 490-495,
 628-633, 634-639, 655-660, 659-664, 852-857, 915-920, 948-953, 987-992

Amidation site.

amino acids 525-528

Cell attachment sequence.

amino acids 372-374

Tyrosine specific protein phosphatases active site.

amino acids 943-955

Protein-tyrosine phosphatase.

amino acids 770-1004

FIGURE 42

MQPPPSLCGRALVALVLACGLSRIWGEERGFPDRATPLLQTAEIMTPPTKTLWPKGSNASLARSLAPAEVPGDRT
AGSPPRRTISPPPCQGPIEIKETFKYINTVVVSLVFLVGIIGNSTLLRIIYKNKCMRNGPNILIASLALGDLLHIVID
IPINVYKLLAEDWPFGAEMCKLVFFIQKASVGITVLSLCALSIDRYRAVASWSRIKGIGVPKWTAVEIVLIWVSVV
LAVPEAIGFDIITMDYKGSYLRIKLLHPVQKTAFMQFYKTAKDWWLFSFYFCLPLAITAFFYTLMTCEMLRKKSGMQ
IALNDHLKQRREVAKTVFCLVLVFALCWLPLHLSRILKLTLYNQNDPNRCELLSFLLVLDYIGINMASLNSCINPIA
LYLVSKRFRKNCFKSCLCCWCQSFEKQSLKFKANDHGYDNFRSSSNKYSSS

Transmembrane domains.

amino acids 1-19, 101-121, 137-157, 177-197, 216-236, 275-295, 323-343, 362-382

N-glycosylation sites.

amino acids 59-62, 119-122

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 302-305

Tyrosine kinase phosphorylation site.

amino acids 424-430

N-myristoylation sites.

amino acids 57-62, 115-120, 170-175, 306-311, 371-376

7 transmembrane receptor homology.

amino acids 118-386